

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 22:59:27 ; Search time 136 Seconds  
(without alignments)

481.378 Million cell updates/sec

Title: AAh26175

Perfect score: 776

Sequence: 1 mkfkhvnsarqykdlwnms.....psptstvtgapsnrpv 149.

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*
- 9: geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	651.5	84.0	482	8	ADR10403 Human pro
2	636	82.0	487	9	ADV50585 Human ATP
3	439	56.6	411	5	ABE57020 Mouse isc
4	300	38.7	413	8	ADP80860 Mouse C13
5	260.5	33.6	501	7	ADC18733 Human cyc
6	212	27.3	52	3	AAQ02942 Human sec
7	137	17.7	29	4	AB68325 c-Jun ami
8	137	17.7	29	7	ADF17925 Synthetic
9	137	17.7	29	8	ADQ01349 C-Jun ami
10	110.5	14.2	1483	8	ADQ97206 Mouse can
11	109	14.0	364	5	AAE18363 Streptoco
12	109	14.0	389	5	ABP25889 Streptoco
13	109	14.0	389	5	AAE18359 Streptoco
14	109	14.0	389	8	ADR83905 S. pyogen
15	109	14.0	447	3	AAH14349 Human Zic
16	109	14.0	447	8	ADN04236 Antipsoi
17	109	14.0	447	8	ADQ17360 Human sof
18	109	14.0	1027	7	ABM85538 Mouse pro
19	108	13.9	447	8	ADU18114 Human ost
20	107	13.8	447	5	ABE57042 Mouse isc
21	106.5	13.7	97	4	AAE62172 Zinc fing
22	106.5	13.7	97	5	AAE28179 Zinc fing
23	106.5	13.7	97	9	ADX44497 Zinc fing
24	105.5	13.6	297	2	AAW08955 Chimeric

25	105.5	13.6	303	2	AAW08956	Chimeric
26	105	791	4	ABB67268	Drosophil	
27	105	13.5	845	4	ABB64280	Drosophil
28	103	13.3	393	9	ABE91544	Microbial
29	103	13.3	515	8	ADQ97211	Human can
30	103	13.3	609	4	ABE57871	Drosophil
31	102.5	13.2	496	2	AAW76984	Mouse Egr
32	102.5	13.2	533	1	AAW76984	Mouse Egr
33	102.5	13.2	533	2	AAW76984	Mouse Egr
34	102.5	13.2	533	4	AAW76984	Human zin
35	102.5	13.2	533	5	ABE57356	Mouse lsc
36	102.5	13.2	533	7	ADF44394	DNA bindi
37	102	13.1	409	2	AAW13845	Glucose r
38	102	13.1	752	4	ABB60124	Drosophil
39	102	13.1	1177	5	ABP74029	Candida a
40	101	13.0	415	7	ADC87625	Mouse KKL
41	101	13.0	415	7	ADC87626	Rat KKL
42	101	13.0	415	7	ADC78984	Rat cell
43	101	13.0	415	7	ADC78983	Murine ce
44	101	13.0	415	9	AAE53103	Mouse kid
45	101	13.0	415	9	AAE53105	Rat kidne

## ALIGNMENTS

RESULT 1  
ADR10403  
ID ADR10403 standard; protein; 482 AA.  
XX  
AC ADR10403;  
XX  
DT 04-NOV-2004 (first entry)  
XX  
DE Human protein useful for treating neurological disease Seq 3909.  
XX  
KW human; oligo-capping method; diagnostic marker; gene therapy;  
KW osteoporosis; neurological disease; Alzheimer's disease;  
KW Parkinson's disease; dementia; short memory; cancer;  
KW sense or motor function; emotional reaction; fear response; panic;  
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;  
KW tranquiliser.  
XX  
OS Homo sapiens.  
XX  
PN EPI447413-A2.  
XX  
PD 18-AUG-2004.  
XX  
PF 12-FEB-2004; 2004EP-00003145.  
XX  
PR 14-FEB-2003; 2003JP-00102207.  
PR 09-MAY-2003; 2003JP-00131452.  
PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;  
PI Wakamatsu A, Ishii S, Nagai K, Irie R;  
XX  
WPI; 2004-583265/57.  
XX  
N-PSDB; ADR08447.  
XX  
PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,  
XX Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.  
XX  
CC Claim 1; SEQ ID NO 3909; 2686pp; English.  
XX  
CC This invention relates to novel, isolated full length human cDNA  
XX molecules and the encoded proteins thereof. Specifically, it refers to  
XX cDNA clones obtained by an oligo-capping method, where none of these  
XX clones are identical to any known human mRNAs. The present invention  
XX describes an immunoassay to identify agonists and antagonists, as well as  
XX antibodies, antisense molecules and siRNAs that can all be used to bind

Best Available Copy

CC to and modulate expression of the cDNA molecules. As such, these  
CC molecules are useful for diagnostic markers or therapeutic targets for  
CC the various diseases or morbid states. In particular, they are useful in  
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's  
CC disease, Parkinson's disease, dementia, short memory and various cancers,  
CC as well as for maintaining equilibrium of sense or motor function, and  
CC for treating emotional reaction, fear response and panic. Accordingly,  
CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,  
CC cytostatic and tranquilliser activities. This polypeptide is a protein  
CC encoded by a full length human cDNA sequence of the invention. NOTE: This  
CC sequence is not given in the sequence listing of the specification but  
CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-  
CC office.  
XX  
SQ Sequence 482 AA;  
Query Match 84.0%; Score 651.5; DB 8; Length 482;  
Best Local Similarity 72.0%; Pred. No. 3.4e-56;  
Matches 134; Conservative 4; Mismatches 11; Indels 37; Gaps 1;  
QY 1 MKFKLVNSARQYKDLWNMSDDKPFCTAGCGQRTNEDHLA----- 43  
DB 1 MKFKLVNSARQYKDLWNMSDDKPFCTAGCGQRTNEDHLA----- 60  
QY 44 -----VHKHKHMTLKFPGARNSKIERSVVTTHQDSPLPHPS 83  
DB 61 LFNELASPFNEFKKASEDDIKMPLDLSPLATPIIRSKIERSVVTTHQDSPLPHPS 120  
QY 84 TTSDEKEVFLAQTAQTSALVRPASIQVNPVLLTSSDSSVVIQQAVPSPSTSTVI 143  
DB 121 TTSDEKEVFLAQTAQTSALVRPASIQVNPVLLTSSDSSVVIQQAVPSPSTSTVI 180  
QY 144 SNRPV 149  
DB 181 SNRPV 186  
RESULT 2  
ADV50585  
ID ADV50585 standard; protein; 487 AA.  
XX  
AC ADV50585;  
DT 24-FEB-2005 (first entry)  
XX  
DE Human ATF-2 protein.  
XX  
KW protein structure; drug screening; ATF-2.  
XX  
OS Homo sapiens.  
XX  
PN WO2004106543-A1.  
XX  
PD 09-DEC-2004.  
XX  
PF 31-MAY-2004; 2004WO-AU000723.  
XX  
PR 30-MAY-2003; 2003US-0474465P.  
XX  
PA (PHYL-) PHYLOGICA LTD.  
XX  
PI Watt PM, Hopkins R, Bogoyevitch M;  
XX  
XX WPI; 2005-021306/02.  
DR  
PT Identifying a region in a protein of interest, useful in drug screening,  
PT comprises expressing a mutated form of the protein of interest and the  
PT native form of the binding partner protein and of other proteins.  
XX  
PS Claim 25; SEQ ID NO 7; 191pp; English.  
XX  
CC The invention relates to a novel method for identifying a region in a  
CC protein that mediates the ability of the protein to bind to a binding

CC partner protein within a protein complex comprising more than two  
CC proteins. The method comprises expressing a mutated form of the protein  
CC of interest and the native form of the binding partner protein and native  
CC forms of one or more other proteins that bind to the protein of interest.  
CC The method of the invention may be useful for identifying a region in a  
CC protein that mediates the ability of the protein to bind to a binding  
CC partner protein within a protein complex. The processes may be useful in  
CC drug screening, for determining or validating a protein interaction as a  
CC therapeutic drug target or validation reagent and for identifying a  
CC therapeutic or prophylactic compound. The current sequence is that of the  
CC human ATF-2 protein of the invention.  
XX  
SQ Sequence 487 AA;  
Query Match 82.0%; Score 636; DB 9; Length 487;  
Best Local Similarity 68.6%; Pred. No. 1.2e-54;  
Matches 131; Conservative 0; Mismatches 0; Indels 60; Gaps 1;  
QY 19 MSDDKPFCTAGCGQRTNEDHLAVHKHMTLKFPGARN----- 60  
DB 1 MSDDKPFCTAGCGQRTNEDHLAVHKHMTLKFPGARNDSVIVADQTPTRFLKN 60  
QY 61 -----SKIEERSVVTTHQDSPL 78  
DB 61 CEEVGLFNLASPFNEFKKASEDDIKMPLDLSPLATPIIRSKIERSVVTTHQDSPL 120  
QY 79 PHESTTTSDEKEVFLAQTAQTSALVRPASIQVNPVLLTSSDSSVVIQQAVPSPSTSTVI 138  
DB 131 PHESTTTSDEKEVFLAQTAQTSALVRPASIQVNPVLLTSSDSSVVIQQAVPSPSTSTVI 180  
QY 139 TQAPSSNRPIV 149  
DB 181 TQAPSSNRPIV 191  
RESULT 3  
ABB57020  
ID ABB57020 standard; protein; 411 AA.  
XX  
AC ABB57020;  
DT 07-MAR-2002 (first entry)  
XX  
DE Mouse ischaemic condition related protein sequence SEQ ID NO:2.  
XX  
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;  
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.  
XX  
OS Mus musculus.  
XX  
PN WO200188188-A2.  
XX  
PD 22-NOV-2001.  
XX  
PF 18-MAY-2001; 2001WO-JP004192.  
XX  
PR 18-MAY-2000; 2000JP-00145977.  
XX  
PA (UTNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.  
XX  
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;  
XX  
XX WPI; 2002-034733/04.  
DR N-PSDB; ABI9202.  
XX  
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring  
PT expression levels of particular genes defined in the specification or by  
PT determining the expression profile of a gene group comprising these  
PT genes.  
XX  
PS Claim 2; Page 51-53; 2690pp; English.  
XX  
CC The present invention describes a method for examining ischaemic

CC conditions, comprising measuring the expression levels of particular  
 CC genes (I) in a test sample or determining the expression profile of a  
 CC gene group in the sample comprising genes selected from (I). The method  
 CC is useful for examining the ischaemic condition (e.g. compressive  
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring the  
 CC expression levels of particular genes (ABI99202 to ABI99912, encoding the  
 CC protein sequences in ABB57020 to ABB57374) or by determining the  
 CC expression profile of a gene group comprising these genes. The expression  
 CC levels or expression profiles produced by these genes are used as an  
 CC indicator when screening for ischaemic condition-improving drugs or  
 CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR  
 CC primers for a mouse ischaemic condition related sequence, which are used  
 CC in the exemplification of the present invention  
 XX  
 SQ Sequence 411 AA;

Query Match 56.6%; Score 439; DB 5; Length 411;  
 Best Local Similarity 95.7%; Pred. No. 4.8e-35;  
 Matches 89; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 57 PARNKIEEVSVEVTHQDSPLPHPESTTSDEKEVPLAQTAQTSATVRPASLOVQPNVLL 116  
 DB 23 PIRSKIEEVSVEVTHQDSPLPHPESTTSDEKEIPLAQTAQTSATVRPASLOVQPNVLL 82  
 QY 117 TSSDSSVIIQAVPSPTSSVITQAPSSNRPIV 149  
 DB 83 TSSDSSVIIQAVPSPTSSVITQAPSSNRPIV 115

RESULT 4  
 ADP80860  
 ID ADP80860 standard; protein; 413 AA.  
 AC ADP80860;  
 DT 09-SEP-2004 (first entry)  
 XX Mouse C130020M04Rik amino acid sequence SEQ ID NO:85.  
 DE  
 XX c-fos; c-fos interacting protein; fos interacting protein chromosome X;  
 KW Fip-cx; screening; Fc-fos interacting protein inhibitor; mouse.  
 XX  
 OS Mus musculus.  
 PN WO2004051121-A1.  
 XX  
 PD 24-JUN-2004.  
 XX  
 PF 19-NOV-2003; 2003WO-JP014749.  
 XX  
 PR 11-DEC-2002; 2002JP-00360046.  
 XX  
 PA (UYKE-) UNIV KEIO.  
 XX  
 PI Miyamoto E, Ishizaka M, Yanagawa H;  
 XX  
 DR WPI; 2004-517250/49.  
 DR N-PSDB; ADP80924.  
 XX  
 XX New proteins that interact with fos, e.g., fos interacting protein  
 PT chromosome X (Fip-cx).  
 XX  
 PS Claim 57; SEQ ID NO 85; 192pp; Japanese.

XX The present invention describes a protein (I) that interacts with c-fos  
 CC (e.g., fos interacting protein chromosome X (Fip-cx), Fip-cx.1, fos  
 CC interacting protein chromosome ex.2, or fos interacting protein  
 CC chromosome 4). Also described: (1) a nucleic acid (II) that encodes (I);  
 CC (2) an inhibitor (III) that inhibits the interaction of (I) encoded by  
 CC (II) with c-fos protein; (3) an inhibitor (IV) that inhibits interaction  
 CC of a protein with c-fos protein; and (4) detecting (M1) the interaction  
 CC of a protein as mentioned in (3) with c-fos protein, by contacting the  
 CC protein with c-fos, to form a composite. (I) is useful for detecting its

CC interaction with c-fos, to form a complex. (M1) is useful for screening  
 CC the protein that interacts with c-fos which involves performing the  
 CC detection process and selecting the protein that interacts with c-fos.  
 CC (M1) is useful for screening a protein that interacts with c-fos which  
 CC involves performing (M1) and selecting the detected protein. (I) is  
 CC useful for screening inhibitors that interact with c-fos. The present  
 CC sequence represents a mouse C130020M04Rik amino acid sequence, which can  
 CC interact with c-fos in the exemplification of the present invention.  
 XX  
 SQ Sequence 413 AA;

Query Match 38.7%; Score 300; DB 8; Length 413;  
 Best Local Similarity 38.8%; Pred. No. 3.7e-21;  
 Matches 76; Conservative 11; Mismatches 33; Indels 76; Gaps 6;  
 QY 19 MSDDKPLCTAPGCGQRTNEDHLAVHKHGHETLKFGPARNSKI----- 63  
 DB 1 MGDPRFVCSAPGCGQRTNEDHLAVHKHGHETLKFGPARTDSVIIATQPTPTFRFLKN 60  
 QY 64 -----EPPSVVETTHQD 75  
 DB 61 CEVGLFNEGLASSFEHFKASDDDEKKGAGPLDMSLPSTPDIKIKEEPVEVDSPPD 120  
 QY 76 SPLPHPESTTSDEKEV---PLAQTAQTSATVRPASLOVQPNVLLTSSDSSVIIQAVPSP 132  
 DB 121 SPASSPCSPPLKEKEVTKFVV-ISTPTTIVRFGSLP----LHLGYDP---LHPTLPSP 172  
 QY 133 TSSTVITQAPSSNRPI 148  
 DB 173 TS--VITOAPSSNRQI 186

RESULT 5  
 ADC18733  
 ID ADC18733 standard; protein; 501 AA.  
 XX  
 AC ADC18733;  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Human cyclic AMP response-element binding protein, CREBPA.  
 XX  
 KW candidate cell death pathway modulating agent;  
 KW cyclic AMP response-element binding protein; CREBPA; cytostatic;  
 KW angiogenic; apoptotic; cell proliferation disorder; cancer; human; gene;  
 KW ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003074672-A2.  
 XX  
 PD 12-SEP-2003.  
 XX  
 PF 28-FEB-2003; 2003WO-US006361.  
 XX  
 PR 01-MAR-2002; 2002US-0360810P.  
 XX  
 PA (EXEL-) EXELIXIS INC.  
 XX  
 PI Costa MA, Parry D, Chen C;  
 XX WPI; 2003-712887/67.  
 DR N-PSDB; ADC18732.  
 XX  
 XX Identifying candidate cell death pathway modulators for treating cell  
 PT death-related diseases, comprises contacting an assay system comprising  
 PT cyclic AMP response-element binding protein, and detecting test agent  
 PT biased activity.  
 XX  
 PS Example I; SEQ ID NO 2; 49pp; English.  
 XX  
 XX The invention relates to a novel method for identifying a candidate cell  
 CC death pathway modulating agent. The novel method comprises contacting an







```
XX SQ Sequence 29 AA;
Query Match 17.7%; Score 137; DB 8; Length 29;
Best Local Similarity 88.9%; Pred. No. 2.2e-06;
Matches 24; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 37 TNEOHLAVHKHGHENTLKFGPARNSKI 63
Db 1 TNEOHLAVHKHGHENTLKFGPARNSDV 27

RESULT 10
ADQ97206
ID ADQ97206 standard; protein; 1483 AA.
XX AC
XX ADQ97206;
XX 07-OCT-2004 (first entry)
XX Mouse cancer associated sequence MP2-08-013, SEQ ID 182.
XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse.
XX OS Mus musculus.
XX WO2004060304-A2.
XX 22-JUL-2004.
XX 22-DEC-2003; 2003WO-US041389.
XX 27-DEC-2002; 2002US-00330773.
XX (SAGR-) SAGRES DISCOVERY INC.
XX Morris DW, Malandro MS;
XX WPI; 2004-543781/52.
XX New isolated cancer associated nucleic acids comprising at least 10
XX contiguous nucleotides, useful for diagnosing, preventing and/or treating
XX cancers such as leukemia and lymphoma.
XX Claim 1; SEQ ID NO 182; 199pp; English.
XX The present invention relates to cancer associated sequences (ADQ97025-
XX ADQ98004). The sequences are useful for the diagnosis, prevention and/or
XX treatment of cancer, such as leukemia and lymphoma. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 1483 AA;
Query Match 14.2%; Score 110.5; DB 8; Length 1483;
Best Local Similarity 26.2%; Pred. No. 0.17; 64; Indels 49; Gaps 9;
Matches 48; Conservative 22; Mismatches 22;

QY 9 SARQYKD---LWNMS-DDKFLCTAGGCGQRTNED----HLAVHKH-----KH 49
Db 649 SATNYKNHVRHITVSPGKPYCTVPGCKRTEYSSLYKHHVHTHCKPYTCSSCGKTY 708
QY 50 EMTLKFGPARNSKIEPSSVETHQ-----DSPLPHPB--STTSDEKEVP 92
Db 709 ROTSTLAWHKRSANGELEATEESEQALYEQQLEAASAAEPPPKPTHIAIYLSVKK-- 766
QY 93 LAQTAQPSAIVRPASLQVNPVLLTSSDSSVLIQAVESP-----TSSTVITQAPSSN 145
Db 767 -ESSAIPTOQAVMTBEDGPPQVALLTDQGT---QQVSUSPEDLQALGSAISVVTORGST 822
QY 146 RPI 148
```

```
Db 823 LTI 825

RESULT 11
AAE18363
ID AAE18363 standard; protein; 364 AA.
XX AC AAE18363;
XX 29-AUG-2002 (revised)
XX 07-MAY-2002 (first entry)
XX Streptococcus pyogenes strain SPY57 BVH-P1 mature protein.
XX BVH-P1 gene; streptococcal infection; pharyngitis; erysipelas; impetigo;
XX scarlet fever; bacteraemia; necrotising fasciitis; toxic shock; vaccine;
XX immune response; anti-inflammatory; immunisation; antibacterial.
XX Streptococcus pyogenes; SPY57.
XX WO2002044495-A2.
XX 17-JAN-2002.
XX 06-JUL-2001; 2001WO-CA001001.
XX 06-JUL-2000; 2000US-0216465P.
XX (SHIR-) SHIRE BIOCHEM INC.
XX Martin D, Hamel J, Brodeur B;
XX WPI; 2002-171701/22.
XX N-PSDB; RAD29298.
XX New Streptococcus pyogenes antigen useful for diagnosing, preventing or
XX treating streptococcal infection, e.g. pharyngitis, erysipelas and
XX impetigo, scarlet fever, and invasive diseases.
XX Claim 21; Fig 14; 74pp; English.
XX The invention relates to antigens, more particularly an antigen of
XX Streptococcus pyogenes (also called group A Streptococcus (GAS))
XX bacterial pathogen. The polypeptides and polynucleotides encoding them
XX are useful for diagnosing, preventing or treating streptococcal
XX infection, such as pharyngitis, erysipelas, impetigo, scarlet fever,
XX invasive diseases (bacteraemia, necrotising fasciitis, toxic shock), and
XX for eliciting an immune response. The polypeptides may also be used as
XX immunogens for producing antibodies for the diagnosis and treatment of
XX Streptococcus infection, or for passive immunisation. DNAs encoding the
XX polypeptides may also be used to design DNA probes for detecting the
XX presence of Streptococcus in biological samples suspected of containing
XX the bacteria. The vaccine composition is useful as a prophylactic or
XX therapeutic treatment of Streptococcal infection in an individual
XX susceptible to or infected with streptococcal infection. The present
XX sequence is Streptococcus pyogenes strain SPY57 BVH-P1 mature protein.
XX (Updated on 29-AUG-2003 to standardise OS field)
XX SQ Sequence 364 AA;
Query Match 14.0%; Score 109; DB 5; Length 364;
Best Local Similarity 28.5%; Pred. No. 0.038;
Matches 35; Conservative 22; Mismatches 36; Indels 30; Gaps 5;

QY 56 GPARNSKIEEP--SVVETHQDSPLPHESTTSDEKEVPLACTAQTPTS-----AIVRPA 107
Db 72 GQATNLTVQAPASSPASVSHVPSSEPLPQASATSOQTVPMAPPATPSDVPTTTPFASAKPD 131
QY 108 SLQVNPVLLTSSDSSV-----ITQQAVPSPSTSS-----TVITOAP-SSN 145
Db 132 SSVTASSELSTNDVSTELSSSKQKQPEVQEAIVTPKAAETTEVEPKTDISEAPTAN 191
QY 146 RPI 148
```

Db 192 RPV 194  
||:  
RESULT 12  
ID ABP25889 standard; protein; 389 AA.  
XX AC ABP25889;  
XX DT 02-JUL-2002 (first entry)  
XX DE Streptococcus polypeptide SEQ ID NO 954.  
XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
XX KW antiinflammatory; infection; vaccine; meningitis; gene therapy.  
XX OS Streptococcus pyogenes.  
XX PN WO200234771-A2.  
XX PD 02-MAY-2002.  
XX PF 29-OCT-2001; 2001WO-GB004789.  
XX PR 27-OCT-2000; 2000GB-00026333.  
XX PR 24-NOV-2000; 2000GB-00028727.  
XX PR 07-MAR-2001; 2001GB-00005640.  
XX XX (CHIR-) CHIRON SPA.  
XX PA (GENO-) INST GENOMIC RES.  
XX PI Telford J, Massignani V, Margarit Y RosI, Grandi G, Fraser C;  
PI Tettelin H;  
XX DR WPI; 2002-352536/38.  
XX DR N-PSDB; ABN66520.  
XX PT New Streptococcus protein for the treatment or prevention of infection or  
XX PT disease caused by Streptococcus bacteria, such as meningitis, and for  
XX PT detecting a compound that binds to the protein.  
XX PS Claim 1; Page 3249; 4525pp; English.  
XX CC The invention relates to a protein (ABP25413-ABP30895) from group B  
XX CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
XX CC the specification. The proteins have antibacterial and antiinflammatory  
XX CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
XX CC antibodies that bind (I) are used in the manufacture of medicaments for  
XX CC the treatment or prevention of infection or disease caused by  
XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
XX CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
XX CC biological sample. (I) is used to determine whether a compound binds to  
XX CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
XX CC used as a vaccine or diagnostic composition. The disease caused by  
XX CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
XX CC acid encoding (I) may be used to recombinantly produce (I) and may be  
XX CC used in gene therapy. Antibodies to (I) are used for affinity  
XX CC chromatography, immunoassays, and distinguishing/identifying  
XX CC Streptococcus proteins  
XX SQ Sequence 389 AA;  
Query Match 14.0%; Score 109; DB 5; Length 389;  
Best Local Similarity 28.5%; Pred. No. 0.042;  
Matches 35; Conservative 22; Mismatches 36; Indels 30; Gaps 5;  
Qy 56 GPARNKIEEP--SVVETHQDSPLPHPESTTSDEKVELPQAQPTS-----AIVRPA 107  
Db 97 GOATNLTVQAPASSPASVSHVPSSEPLFQASATSOPTVPMAPPATPSDVPTTFFASAKPD 156

Qy 108 SLOQPNVLLTSSDSV-----IIQAAVPSPTSS-----TVITQAP-SSN 145  
Db 157 SSVTASSSLTSTNDVSTELSSSQKQPEVQAEVPTPKAAETTEVEPKTDISEAPTSAN 216  
Qy 146 RPI 148  
Db 217 RPV 219  
RESULT 13  
AAE18359  
ID AAE18359 standard; protein; 389 AA.  
XX AC AAE18359;  
XX DT 29-AUG-2003 (revised)  
XX DT 07-MAY-2002 (first entry)  
XX DE Streptococcus pyogenes strain SPY57 BVH-P1 protein.  
XX KW BVH-P1 gene; streptococcal infection; pharyngitis; erysipelas; impetigo;  
XX KW scarlet fever; bacteraemia; necrotising fascitis; toxic shock; vaccine;  
XX KW immune response; anti-inflammatory; immunisation; antibacterial.  
XX OS Streptococcus pyogenes; SPY57.  
XX FH Key Location/Qualifiers  
XX FT Peptide 1..25  
XX FT /label= Signal\_peptide  
XX FT Protein 26..389  
XX FT /note= "Mature\_BVH\_P1\_protein"  
XX PN WO200204495-A2.  
XX PD 17-JAN-2002.  
XX PF 06-JUL-2001; 2001WO-CA001001.  
XX PR 06-JUL-2000; 2000US-0216465P.  
XX XX (SHIR-) SHIRE BIOCHEM INC.  
XX PI Martin D, Hamel J, Brodeur B;  
XX DR WPI; 2002-171701/22.  
XX DR N-PSDB; AAD28294.  
XX PT New Streptococcus pyogenes antigen useful for diagnosing, preventing or  
XX PT treating streptococcal infection, e.g. pharyngitis, erysipelas and  
XX PT impetigo, scarlet fever, and invasive diseases.  
XX PS Claim 21; Fig 6; 74pp; English.  
XX CC The invention relates to antigens, more particularly an antigen of  
XX CC Streptococcus pyogenes (also called group A Streptococcus (GAS))  
XX CC bacterial pathogen. The polypeptides and polynucleotides encoding them  
XX CC are useful for diagnosing, preventing or treating streptococcal  
XX CC infection, such as pharyngitis, erysipelas, impetigo, scarlet fever,  
XX CC invasive diseases (bacteraemia, necrotising fascitis, toxic shock), and  
XX CC for eliciting an immune response. The polypeptides may also be used as  
XX CC immunogens for producing antibodies for the diagnosis and treatment of  
XX CC Streptococcus infection, or for passive immunisation. DNAs encoding  
XX CC polypeptides may also be used to design DNA probes for detecting the  
XX CC presence of Streptococcus in biological samples suspected of containing  
XX CC the bacteria. The vaccine composition is useful as a prophylactic or  
XX CC therapeutic treatment of streptococcal infection in an individual  
XX CC susceptible to or infected with streptococcal infection. The present  
XX CC sequence is Streptococcus pyogenes strain SPY57 BVH-P1 protein. (Updated  
XX CC on 29-AUG-2003 to standardise OS field)  
XX SQ Sequence 389 AA;  
Query Match 14.0%; Score 109; DB 5; Length 389;

Best Local Similarity 28.5%; Pred. No. 0.042;  
Matches 35; Conservative 22; Mismatches 36; Indels 30; Gaps 5;

QY 56 GPARNSKIEEP--SVVETTHQDSPLPHPPESTTSDEKEVPLAQTAQPTS-----AIVRPA 107  
DB 97 GQATNLTVQAPASSPASVSHVPSSEPLQASATSQPTVPMAPPATPSDVPTTFFASAKPD 156  
QY 108 SLOVPNVLLTSSDSSV-----IIQNAVPSPTSS-----TVITQAP-SSN 145  
DB 157 SSVTASSELSTSTNDVSTELSESQKQPEVQEAAPTTPKAAETTEVEPKTDISEAPTGSAN 216  
QY 146 RPI 148  
DB 217 RPV 219

RESULT 14  
ADR83905  
ID ADR83905 standard; protein; 389 AA.  
AC ADR83905;  
DT 02-DEC-2004 (first entry)  
DE S. pyogenes hyperimmune system reactive antigen Spy0469.  
XX hyperimmune serum reactive antigen; vaccine; anticaline.  
KW Streptococcus pyogenes.  
OS  
PN WO2004078907-A2.  
XX  
PD 16-SEP-2004.  
XX  
PF 02-MAR-2004; 2004WO-BP002087.  
XX  
PR 04-MAR-2003; 2003EP-00450061.  
XX  
XX (INTE-) INTERCELL AG.  
PA  
PI Meinke A, Nagy E, Winkler B, Gelbmann D;  
XX  
XX WPI; 2004-653698/63.  
DR N-PSDB; ADR83755.  
XX

New isolated nucleic acid molecules encoding hyperimmune serum-reactive antigens from Streptococcus pyogenes, useful for diagnosing, preventing and treating S. pyogenes infections.

Claim 14; SEQ ID NO 173; 145pp; English.

This invention describes a novel nucleic acid molecule encoding a hyperimmune serum reactive antigen or its fragment from Streptococcus pyogenes. The nucleic acid molecule or hyperimmune serum-reactive antigen or its fragment are useful for the manufacture of a pharmaceutical preparation, especially a vaccine, against S. pyogenes infection. In addition, the hyperimmune serum reactive antigen or fragment is used for the isolation and/or purification and/or identification of an interaction partner of the hyperimmune serum reactive antigen or its fragment, for the generation of a peptide (e.g. anticalines) binding to the antigen or fragment, or for the manufacture of a functional nucleic acid selected from aptamers and Spiegelmers. The nucleic acid molecule may also be used for the manufacture of functional ribonucleic acids, such as ribozymes, antisense nucleic acids and siRNA. ADR83733-ADR84189 represent S. pyogenes hyperimmune serum reactive antigens, fragments and the encoding polynucleotide described in the invention.

Query Match  
Best Local Similarity 28.5%; Pred. No. 0.042;  
Matches 35; Conservative 22; Mismatches 36; Indels 30; Gaps 5;

QY 56 GPARNSKIEEP--SVVETTHQDSPLPHPPESTTSDEKEVPLAQTAQPTS-----AIVRPA 107  
DB 97 GQATNLTVQAPASSPASVSHVPSSEPLQASATSQPTVPMAPPATPSDVPTTFFASAKPD 156  
QY 108 SLOVPNVLLTSSDSSV-----IIQNAVPSPTSS-----TVITQAP-SSN 145  
DB 157 SSVTASSELSTSTNDVSTELSESQKQPEVQEAAPTTPKAAETTEVEPKTDISEAPTGSAN 216  
QY 146 RPI 148  
DB 217 RPV 219

RESULT 15  
AAB14349  
ID AAB14349 standard; protein; 447 AA.  
XX AAB14349;  
AC  
DT 29-NOV-2000 (first entry)  
XX  
DE Human Zic1 protein.  
XX  
XX Human; Zic1; epidermal cell transdifferentiation; gene therapy;  
KW cerebroprotective; neuroprotective; brain injury; spinal cord injury;  
KW stroke; neurodegenerative disease; Parkinson's disease;  
KW Huntington's disease; Alzheimer's disease; neuronal cell generation.  
XX  
OS Homo sapiens.  
XX  
XX US6087168-A.  
XX  
PD 11-JUL-2000.  
XX  
PF 20-JAN-1999; 99US-00234332.  
XX  
PR 20-JAN-1999; 99US-00234332.  
XX  
XX (CEDA-) CEDARS SINAI MEDICAL CENT.  
PA  
PI Levesque MF, Neuman T;  
XX  
XX WPI; 2000-498200/44.  
DR N-PSDB; AAA62683.  
XX

Converting epidermal cells into neurons, useful for isolating nerve growth factors or for gene therapy, comprises dedifferentiating cells and transfecting with vectors with a cDNA coding neurogenic transcription factors.

Example 2; Col 37-40; 27pp; English.

The present sequence is the Zic1 protein, which is encoded by a human gene from Genbank. The Genbank sequence was used to provide sequence information for the cloning of Zic1 cDNA, which was used to transfect cultured epidermal cells. This was part of a novel method for transdifferentiating an epidermal basal cell into a cell having the morphological, physiological and/or immunological features of a viable neuronal cell. The method is useful for screening new drugs for treating a nervous system disorder, or for isolating a novel nerve growth factor. The transdifferentiated cell is useful in both cell and gene therapies aimed at alleviating various neurological disorders. The cell or gene therapy approach involves the use of autologous transplantation or grafting of the newly created neuronal cells as treatment for brain or spinal cord injury, stroke and neurodegenerative diseases (e.g. Parkinson's disease, Huntington's disease or Alzheimer's disease)

Query Match  
Best Local Similarity 26.1%; Pred. No. 0.05;  
Matches 42; Conservative 25; Mismatches 58; Indels 36; Gaps 7;

```
QY      2  KFKLHVNSARQYKDLWNMSDDKPFCLCTAPCGCORFTNEDHLAVHKHKHEMTLKFGP----- 57
Db      286 KYKL-VNHIRVH-----TGEKPPCPFPFGCGKVFAENSENLKHKRTHTGKFPFKCEPEG 338
QY      58  -----ARNSKIEEPSVVETTHQ-----DSPLPHPESTTSDEK-EVPLAQTAQPTSAIV 104
Db      339 CDRRFANSDDRKKHMHVHTSDKPYLCKMCDKSYTHPSSVVRKHKVHSSSSQGSQSPSPA-- 396
QY      105 RPASLQVNPVLLTSSDSSVIIQQAVPSPTSSTVITQAPSSN 145
Db      397 -----ASSGYESTPPTIVSPSTDNPTSSL-----SPSSS 427
```

Search completed: January 10, 2006, 23:05:08  
Job time : 138 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 23:01:12 ; Search time 17 Seconds  
(without alignments)  
843.312 Million cell updates/sec

Title: AAH26175  
Perfect score: 776  
Sequence: 1 mkfklhvnсарqkdlwnms.....psptestvitqapsnrpiv 149

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	736	94.8	505	1 S05380	transcription fact
2	594	76.5	486	2 J04028	activating transcr
3	450	58.0	456	2 B42026	cyclic AMP respo
4	445	57.3	448	2 A42026	CAMP response elem
5	363	46.8	389	1 A39429	CAMP response elem
6	295	38.0	483	2 S12741	transcription fact
7	260.5	33.6	508	2 A45477	CAMP response elem
8	180	23.2	358	2 C42026	cyclic AMP respo
9	170	21.9	313	2 A34785	DNA-binding protei
10	113	14.6	428	2 T43532	zinc finger protei
11	107	13.8	447	2 I56511	Zic protein - mous
12	107	13.8	515	2 A44256	zinc-finger protei
13	107	13.8	582	2 T38743	hypothetical zinc
14	103	13.3	609	2 A49839	odd-paired - fruit
15	102.5	13.2	533	2 J50304	developmental cont
16	102.5	13.2	1367	1 S48478	glucan 1,4-alpha-g
17	102	13.1	409	2 S70704	carbon catabolite
18	102	13.1	565	2 T39863	zinc finger protei
19	100.5	13.0	1361	2 T30864	neural specific DN
20	99.5	12.8	593	2 J07829	metal-responsive t
21	98	12.6	532	2 C27993	protein C27A12.3 [
22	97.5	12.6	696	2 A29635	transcription fact
23	97	12.5	1402	2 S42748	finger protein - f
24	95	12.2	353	2 J04875	zinc-finger protei
25	95	12.2	697	2 B44489	GT box-binding pro
26	95	12.2	1350	2 T30341	zinc finger protei
27	94.5	12.2	1323	2 T30253	spalt protein - mo
28	94.5	12.2	1360	2 T12064	DNA binding protei
29	94	12.1	392	2 S11998	finger protein odd

30	93.5	12.0	506	2 A40679	transcription enha
31	93.5	12.0	523	2 B40679	transcription enha
32	93.5	12.0	1596	2 A35927	190K DNA-binding p
33	93	12.0	264	2 T27830	hypothetical prote
34	93	12.0	402	2 S70703	carbon catabolite
35	93	12.0	770	1 TWBYA2	transcription fact
36	92.5	11.9	788	2 J50747	regulatory protein
37	92.5	11.9	914	2 S46593	finger protein AZF
38	92	11.9	564	2 I53106	gene gli protein -
39	91.5	11.8	543	2 A41211	early growth respo
40	91	11.7	479	2 T46318	hypothetical prote
41	91	11.7	615	2 T34392	hypothetical prote
42	90.5	11.7	292	2 S24169	mucin - rat
43	90.5	11.7	457	2 I54340	DNA-binding protei
44	90.5	11.7	688	2 A56360	zinc finger protei
45	90	11.6	508	2 A32225	nerve growth facto

## ALIGNMENTS

### RESULT 1

S05380  
transcription factor ATF2 - human  
N:Alternate names: activating transcription factor 2 (ATF-2); cAMP response element-bin  
N:Contains: CAMP response element-binding protein HB16  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: S05380; A34776; B34223  
R:Maekawa, T.; Sakura, H.; Kanei-Ishii, C.; Sudo, T.; Yoshimura, T.; Fujisawa, J.I.; Yo  
EMBO J. 8, 2023-2028, 1989  
A:Title: Leucine zipper structure of the protein CRE-BP1 binding to the cyclic AMP resp  
A:Reference number: S05380; MUID:90005408; PMID:2529117  
A:Accession: S05380  
A:Molecule type: mRNA  
A:Residues: 1-505 <MAE>  
A:Cross-references: UNIPROT:P15336; UNIPARC:UPI000016A741; EMBL:X15875; NID:G30214; PID  
R:Kara, C.J.; Liou, H.C.; Ivashkiv, L.B.; Glimcher, L.H.  
Mol. Cell. Biol. 10, 1347-1357, 1990  
A:Title: A cDNA for a human cyclic AMP response element-binding protein which is distin  
A:Reference number: A34776; MUID:90205810; PMID:2320002  
A:Accession: A34776  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 211-222, N', 224-505 <XAR>  
A:Cross-references: UNIPARC:UPI000016AA16; GB:M31630; NID:G183787; PIDN:AAA35951.1; PID  
R:Hal, T.; Liu, F.; Coukos, W.J.; Green, M.R.  
Genes Dev. 3, 2083-2090, 1989  
A:Title: Transcription factor ATF cDNA clones: an extensive family of leucine zipper pr  
A:Reference number: A91622; MUID:90185187; PMID:2516827  
A:Accession: B34223  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 107-357, 'V', 359-465 <HA2>  
A:Cross-references: UNIPARC:UPI00001748C4  
C:Genetics:  
A:Gene: GDB:ATF2; CREB2; TREB7; CRE-BP1  
A:Cross-references: GDB:128011; OMIM:123811  
A:Map position: 2q32-2q32  
C:Superfamily: CAMP response element-binding protein 1; fos/jun DNA-binding domain homo  
C:Keywords: alternative splicing; DNA binding; nucleus; transcription regulation  
F/347-387/domain: fos/jun DNA-binding domain homology <FJD>

Query Match	94.8%	Score	736	DB	1	Length	505
Best Local Similarity	71.3%	Pred. No.	7.9e-55				
Matches	149	Conservative	0	Mismatches	60	Gaps	1
Qy	1	MKFKLHVN SARQYKDLWNMSDDKFLCTAPCGGRTNEDHLAVHKHKHMTLKFGPARN	60				
Db	1	MKFKLHVN SARQYKDLWNMSDDKFLCTAPCGGRTNEDHLAVHKHKHMTLKFGPARN	60				
Qy	61	-----	61				

Db 61 DSVIVADQPTPTFLKNCVEVGLFNLASPFENEFKASEDDIKMPLDLSPLATPIIR 120

QY 61 SKIEPSVVTTHQDSPLPHPESTTSDEKEVPLAQTSAIVRPASLOVNVLLTSSD 120

Db 121 SKIEPSVVTTHQDSPLPHPESTTSDEKEVPLAQTSAIVRPASLOVNVLLTSSD 180

QY 121 SSVIIQQAVPSTSSVTITQAPSSNRPIV 149

Db 181 SSVIIQQAVPSTSSVTITQAPSSNRPIV 209

RESULT 2

JC4028

activating transcription factor 2 - African clawed frog

N:Alternate names: cyclic AMP-response element-binding protein

C:Species: Xenopus laevis (African clawed frog)

C>Date: 13-Jun-1995 #sequence\_revision 14-Jul-1995 #text\_change 09-Jul-2004

C:Accession: JC4028

R:Villarreal, X.C.; Richter, J.D.

Gene 153, 225-229, 1995

A:Title: Analysis of ATF2 gene expression during early Xenopus laevis developme nt.

A:Reference number: JC4028; MUID:95180723; PMID:7875593

A:Accession: JC4028

A:Molecule type: mRNA

A:Residues: 1-486 <VIL>

A:Cross-references: UNIPROT:Q91576; UNIPARC:UPI00000FB2FD; GB:U16158; NID:Q887779; PIDN:

C:Comment: This protein is a sequence-specific DNA-binding protein that mediates transac

C:Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homol

C:Keywords: leucine zipper; phosphoprotein; transcription regulation

F:37-398/Region: leucine zipper motif

F:329-369/Domain: fos/jun DNA-binding domain homology <FJD>

F:332-364/Region: basic

F:82/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted

F:102/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 76.5%; Score 594; DB 2; Length 486;

Best Local Similarity 64.9%; Pred. No. 8e-43;

Matches 124; Conservative 4; Mismatches 3; Indels 60; Gaps 2;

QY 19 MSDDKPFCLTAPGCGRFTNEDHLAVHKHKHMTLKGPARN-----SKIEPSVVTTHQDSPLP 79

Db 1 MSDDKPFCLTAPGCGRFTNEDHLAVHKHKHMTLKGPARNDLSVIVADQPTPTFLKX 60

QY 61 -----SKIEPSVVTTHQDSPLP 120

Db 61 CBEVGLFNLASPFENEFKASEDDIKMPLDLSPLATPIIRSKIEPSVVTTHQDSPLP 120

QY 80 HPESITSDS-KEVPLAQTSAIVRPASLOVNVLLTSSDSSSVIIQQAVPSTSSVTI 138

Db 121 HPESITSDHQNVPLAQTSAIVRPASLOVNVLLTSSDSSSVIIQQAVPSTSSVTI 180

QY 139 TQAPSSNRPIV 149

Db 181 TQAPSTNSSIV 191

RESULT 3

B42026

cyclic AMP response element DNA-binding protein isoform 1 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 26-Aug-1999

C:Accession: B42026

R:Georgopoulos, K.; Morgan, B.A.; Moore, D.D.

Mol. Cell. Biol. 12, 747-757, 1992

A:Title: Functionally distinct isoforms of the CRE-BP DNA-binding protein mediate activi

A:Reference number: A42026; MUID:92123199; PMID:1531087

A:Contents: EL4

A:Accession: B42026

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-456 <GEO>

A:Cross-references: UNIPARC:UPI0000170CB6; GB:S76657; NID:G243428; PIDN:AAB21128.1; PID:

A>Note: sequence extracted from NCBI backbone (NCBIN:76657, NCBIIP:76658)

C:Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homo

F:297-337/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 58.0%; Score 450; DB 2; Length 456;

Best Local Similarity 62.3%; Pred. No. 1.2e-30;

Matches 99; Conservative 0; Mismatches 0; Indels 60; Gaps 1;

QY 51 MTLKFGPARN-----

Db 1 MTLKFGPARNDLSVIVADQPTPTFLKNCVEVGLFNLASPFENEFKASEDDIKMPLD 60

QY 61 -----SKIEPSVVTTHQDSPLPHPESTTSDEKEVPLAQTSAIVRPASLO 110

Db 61 LSPLATPIIRSKIEPSVVTTHQDSPLPHPESTTSDEKEVPLAQTSAIVRPASLO 120

QY 111 VNVLLTSSDSSSVIIQQAVPSTSSVTITQAPSSNRPIV 149

Db 121 VNVLLTSSDSSSVIIQQAVPSTSSVTITQAPSSNRPIV 159

RESULT 4

A42026

cAMP response element-binding protein 3 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 26-Aug-1999

C:Accession: A42026

R:Georgopoulos, K.; Morgan, B.A.; Moore, D.D.

Mol. Cell. Biol. 12, 747-757, 1992

A:Title: Functionally distinct isoforms of the CRE-BP DNA-binding protein mediate activ

A:Reference number: A42026; MUID:92123199; PMID:1531087

A:Contents: EL4

A:Accession: A42026

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-448 <GEO>

A:Cross-references: UNIPARC:UPI0000170CB5; GB:S76655; NID:G243426; PIDN:AAB21127.1; PID:

A>Note: sequence extracted from NCBI backbone (NCBIN:76655, NCBIIP:76656)

C:Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homo

F:289-329/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 57.3%; Score 445; DB 2; Length 448;

Best Local Similarity 70.1%; Pred. No. 3e-30;

Matches 103; Conservative 1; Mismatches 29; Indels 14; Gaps 4;

QY 17 WMMSDKP-----FL---CTAPG---CGQRTNEDHLAVHKHKHMTLKGPAR-----NSK 62

Db 5 WWPQDTPTPTFLKNCVEVGLFNLASPFENEFKASEDDIKMPLDLSPLATPIIRSK 64

QY 63 IEPSVVTTHQDSPLPHPESTTSDEKEVPLAQTSAIVRPASLOVNVLLTSSDSS 122

Db 65 IEPSVVTTHQDSPLPHPESTTSDEKEVPLAQTSAIVRPASLOVNVLLTSSDSS 124

QY 123 VIIQQAVPSTSSVTITQAPSSNRPIV 149

Db 125 VIIQQAVPSTSSVTITQAPSSNRPIV 151

RESULT 5

A39429

cAMP response element-binding protein ATF2 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C:Accession: A39429

R:Kagayana, R.; Sasai, Y.; Nakanishi, S.

J. Biol. Chem. 266, 15225-15231, 1991

A:Title: Molecular characterization of transcription factors that bind to the cAMP resp

A:Reference number: A39429; MUID:91332085; PMID:1714459

A:Accession: A39429

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-389 <KAG>

A:Cross-references: UNIPROT:Q00969; UNIPARC:UPI000002A474; GB:M65148; NID:G206569; PIDN:

C:Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homc



C;Keywords: DNA binding; nucleus; transcription regulation  
F:231-271/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 46.8%; Score 363; DB 1; Length 389;  
Best Local Similarity 44.9%; Pred. No. 2.2e-23;  
Matches 79; Conservative 7; Mismatches 4; Indels 86; Gaps 4;

QY 19 MSDDKPFLLCTAPGCGQRTNEDHLAVHKHHEMTLKGPARN-----SKIEPSVVEVTHQDSPL 60  
DB 1 MSDDKPFLLCTAPGCGQRTNEDHLAVHKHHEMTLKGPARNDSIVADQTPPTTRFLKN 60  
QY 61 -----SKIEPSVVEVTHQDSPL 78  
DB 61 CEEVGLFNLASPFNEFFKASEDDIKMPLDLSPPLATPIIRSKIEPSVVEVTHQDSPL 120  
QY 79 PPESTTSDEKEVPLAQTAQTSIAIVRPASLOVNVLLTSSDSSVVIQAVPSPS 134  
DB 121 PPESTTNDK-----LVRPVTM-VPSV-----PGIPGSS 150

## RESULT 6

transcription factor ATF-a - human  
C;Species: Homo sapiens (man)  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 26-Aug-1999  
C;Accession: S12741  
R;Gaire, M.; Chatton, B.; Keding, C.  
Nucleic Acids Res. 18, 3467-3473, 1990  
A;Title: Isolation and characterization of two novel, closely related ATF cDNA clones from human T-lymphocytes.  
A;Reference number: S12741; MUID:90301459; PMID:1694576  
A;Accession: S12741  
A;Molecule type: mRNA  
A;Residues: 1-483 <GAI>  
A;Cross-references: UNIPARC:UPI0000161877; EMBL:X52943; NID:928912; PIDN:CRA37118.1; PIDN:125836; NCBIP:125837  
C;Superfamily: CAMP response element-binding protein 1; fos/jun DNA-binding domain homology <FJD>  
F:327-367/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 38.0%; Score 295; DB 2; Length 483;  
Best Local Similarity 38.7%; Pred. No. 1.7e-17;  
Matches 75; Conservative 10; Mismatches 33; Indels 76; Gaps 6;

QY 19 MSDDKPFLLCTAPGCGQRTNEDHLAVHKHHEMTLKGPARNSKI----- 63  
DB 1 MGDDRFVVCNAPGCGQRTNEDHLAVHKHHEMTLKGPARNSKIADQTPPTTRFLKN 60  
QY 64 -----SEPSVVEVTHQD 75  
DB 61 CEEVGLFNLASPFNEFFKAADEDEKAAAGPLDMSLPSTPDIIKIEEPVEVDSPPD 120  
QY 76 SPLPPESTTSDEKEV---PLAQTAQTSIAIVRPASLOVNVLLTSSDSSVVIQAVPSP 132  
DB 121 SPASSPCSPPLKEVTPKVL-ISTPTPTIVRPGSLP-----LHLGYDP---LHFTLSP 172

## RESULT 7

CAMP response element-binding protein CRE-BPA - human  
C;Species: Homo sapiens (man)  
C;Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C;Accession: A45477; I58379; I78878  
R;Nomura, N.; Zu, Y.L.; Maekawa, T.; Tabata, S.; Akiyama, T.; Ishii, S.  
J. Biol. Chem. 268, 4259-4266, 1993  
A;Title: Isolation and characterization of a novel member of the gene family encoding the CRE-BPA protein.  
A;Reference number: A45477; MUID:93179432; PMID:8440710  
A;Accession: A45477  
A;Molecule type: mRNA  
A;Residues: 1-508 <NOM>  
A;Cross-references: UNIPROT:Q02930; UNIPARC:UPI00000747EA; GB:L05515; NID:g181049; PIDN:

A;Experimental source: KG-1 cells  
A;Note: sequence extracted from NCBI backbone (NCBIN:125836, NCBIP:125837)  
R;Zu, Y.L.; Maekawa, T.; Nomura, N.; Nakata, T.; Ishii, S.  
Oncogene 8, 2749-2758, 1993  
A;Title: Regulation of trans-activating capacity of CRE-BPA by phorbol ester tumor promoter TPA.  
A;Reference number: I58379; MUID:93390949; PMID:8378084  
A;Accession: I58379  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 8-508 <ZUY1>  
A;Cross-references: UNIPARC:UPI000002A177; GB:L05911; NID:g181051; PIDN:AAC37525.1; PIDN:125836; NCBIP:125837  
A;Accession: I78878  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 34-508 <ZUY2>  
A;Cross-references: UNIPARC:UPI000002A178; GB:L05913; NID:g181055; PIDN:AAC37526.1; PIDN:125836; NCBIP:125837  
C;Genetics: GDB:9957436  
A;Gene: GDB:CRE-BPA  
A;Cross-references: GDB:9957436  
A;Map position: 7p15-7p15  
C;Superfamily: CAMP response element-binding protein 1; fos/jun DNA-binding domain homology <FJD>  
F:370-410/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 33.6%; Score 260.5; DB 2; Length 508;  
Best Local Similarity 37.8%; Pred. No. 1.5e-14;  
Matches 59; Conservative 21; Mismatches 41; Indels 35; Gaps 3;

QY 18 NMSDDKPFLLCTAPGCGQRTNEDHLAVHKHHEMTLKGPARNSKIEPSVVEVTHQDSPL 77  
DB 9 NLEQRPVCSAPGCSQRPFTEDHLMIRHKHEMTLKF-----PSIKTDNLSQ 58  
QY 78 LPHP-----ESTTSDEKEVPLAQTAQPT-----SAIVRPASLOVP 112  
DB 59 TPTPTFLKNCNEVGLFSELDCSLEHPEKKAQEESSKRNISMHNAVGMTPGPGTHOLS 118  
QY 113 NVLLTSSDSSVVIQAVPSPSTSVITQAPSSNRPI 148  
DB 119 SARLPNHDNTNVIQOAMPSSQSSSVITQAPSTNRQI 154

## RESULT 8

Cyclic AMP response element DNA-binding protein isoform 2 - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 26-Aug-1999  
C;Accession: C42026  
R;Georgopoulos, K.; Morgan, B.A.; Moore, D.D.  
Mol. Cell. Biol. 12, 747-757, 1992  
A;Title: Functionally distinct isoforms of the CRE-BP DNA-binding protein mediate activation of the c-fos promoter.  
A;Reference number: A42026; MUID:92123199; PMID:1531087  
A;Contents: EL4  
A;Accession: C42026  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-358 <GEO>  
A;Cross-references: UNIPARC:UPI0000170CB7; GB:S76659; NID:g243430; PIDN:AAB1129.1; PIDN:125836; NCBIP:125837  
A;Note: sequence extracted from NCBI backbone (NCBIN:76659, NCBIP:76660)  
C;Superfamily: CAMP response element-binding protein 1; fos/jun DNA-binding domain homology <FJD>  
F:199-239/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 23.2%; Score 180; DB 2; Length 358;  
Best Local Similarity 33.3%; Pred. No. 6.4e-08;  
Matches 48; Conservative 6; Mismatches 4; Indels 86; Gaps 4;

QY 51 MTLKFGPARN-----SKIEPSVVEVTHQDSPLPHEPSTTSDEKEVPLAQTAQTSIAIVRPASLO 60  
DB 1 MTLKFGPARNSVIVADQTPPTPTPLKNCNEVGLFNLASPFNEFFKASEDDIKMPLD 60  
QY 61 -----SKIEPSVVEVTHQDSPLPHEPSTTSDEKEVPLAQTAQTSIAIVRPASLO 110  
DB 61 LSPLATPIIRSKIEPSVVEVTHQDSPLPHEPSTTSDEK-----LVRPVTM- 106  
QY 111 VPNVLLTSSDSSVVIQAVPSPS 134

Db 107 VFSV-----:|:|:|:|:|:|:PGIPGPSS 118

## RESULT 9

A34785  
DNA-binding protein mXBP - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 13-Jul-1990 #sequence\_revision 13-Jul-1990 #text\_change 26-Aug-1999  
C:Accession: A34785  
R:Ivaahkiv, L.B.; Liou, H.C.; Kara, C.J.; Lamph, W.W.; Verma, I.M.; Glimcher, L.H.  
Mol. Cell. Biol. 10, 1609-1621, 1990  
A:Title: mXBP/CRE-BP2 and c-Jun form a complex which binds to the cyclic AMP, but not to  
A:Reference number: A34785; MUID:90205841; PMID:2138707  
A:Accession: A34785  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-313 <I>A>  
A:Cross-references: UNIPARC:UPI000016CF39; GB:M31629; NID:gl99926; PIDN:AAA39780.1; PID:  
C:Superfamily: CAMP response element-binding protein 1; fos/jun DNA-binding domain homol  
C:Keywords: DNA binding; nucleus; transcription regulation  
F:155-195/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 21.9%; Score 170; DB 2; Length 313;  
Best Local Similarity 50.0%; Pred. No. 3.8e-07;  
Matches 39; Conservative 6; Mismatches 7; Indels 26; Gaps 3;

QY 57 PARNKIEPVSVEVTHQDPLPHEPSTTSDEKEVPLAQTAQTSIAIVRPASLOQVNVLL 116  
DB 23 PIIRSKIEPVSVEVTHQDPLPHEPSTTSDEK-----LVRPVTM-VPSV-- 66

QY 117 TSSDSSVLIQAVPSPTS 134  
DB 67 -----PGIPGPSS 74

## RESULT 10

T43532  
zinc finger protein rev1 - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C>Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 04-Feb-2000  
C:Accession: T43532  
R:Haio, Z.; Furunobu, A.; Nagata, A.; Okayama, H.  
J. Cell Sci. 110, 2557-2566, 1997  
A:Title: A zinc finger protein required for stationary phase viability in fission yeast.  
A:Reference number: 222548; MUID:98039713; PMID:9372444  
A:Accession: T43532  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-428 <HAO>  
A:Cross-references: UNIPARC:UPI00001687F4; EMBL:AF016222; PIDN:AAB87047.1  
A:Experimental source: strain 972h(-)  
A:Gene: rev1  
C:Genetics:  
C:Function:  
A:Description: required for cell viability in a stationary phase induced by glucose star

Query Match 14.6%; Score 113; DB 2; Length 428;  
Best Local Similarity 21.4%; Pred. No. 0.037;  
Matches 39; Conservative 20; Mismatches 65; Indels 58; Gaps 4;

QY 12 QYKDLNMSDDKPLCTAGCGQRTNEDHLAVHKHKTMLKFGPARNKIEEPSVET 71  
DB 19 QVRHRSHTGKPFECSPCKKRTFRDELIRHVRTH-----LRKALVTFEQTLDV 70

QY 72 THQDPLPHEPSTTSDEKEVPLAQT-----AQPTSAIVRPASLI 109  
DB 71 NLHTAPDSKPEGKSTGQEKADKSONSKGSIITDPVQAALVLSVAYAKFTSVLSPTDL 130

QY 110 QVNPVLLT-----SSDSSVLIQAVPSPTSVTITQAPS 144  
DB 131 QAQSKLIEKPRRSASNATGSLNKNQDPLRRFSIESAGAAAPTSPSNS---KSPPE 187

QY 145 NR 146  
DB 188 NR 189

## RESULT 11

I56511  
Zic protein - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I56511  
R:Aruga, J.; Yokota, N.; Hashimoto, M.; Furutachi, T.; Fukuda, M.; Mikoshiba, K.  
J. Neurochem. 63, 1880-1890, 1994  
A:Title: A novel zinc finger protein, zic, is involved in neurogenesis, especially in t  
A:Reference number: I56511; MUID:95016733; PMID:7931345  
A:Accession: I56511  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-447 <RES>  
A:Cross-references: UNIPARC:UPI0000029DC7; GB:D32167; NID:g565654; PIDN  
C:Genetics:  
A:Gene: zic

Query Match 13.8%; Score 107; DB 2; Length 447;  
Best Local Similarity 26.1%; Pred. No. 0.13;  
Matches 42; Conservative 24; Mismatches 59; Indels 36; Gaps 7;

QY 2 KFKLVNSARQYKDLNMSDDKPLCTAGCGQRTNEDHLAVHKHKTMLKFGP---- 57  
DB 286 KYKL-VNHIRVH-----TGEKPPFPFGCGKVFARSENLIKHKRTHTGEKPFKCEFG 338

QY 58 -----ARNSKIEEPSVVEVTHQ-----DSPLPHEPSTTSDEK-EVPLAQTAQTSIAIV 104  
DB 339 CDRFPANSSDRKGMHVHTSDKPYLCKMCKSYTHPSSLRKHKMKVHSSQSGQSPSA-- 396

QY 105 RPASLOQVNPVLLTSSDSSVLIQAVPSPTSVTITQAPSSN 145  
DB 397 -----ASSGYESTPTPTIVSPTTNDNPTTSM---SPSSS 427

## RESULT 12

A44256  
zinc-finger protein ZNF76 - human  
C:Species: Homo sapiens (man)  
C>Date: 30-Apr-1993 #sequence\_revision 03-May-1996 #text\_change 05-Nov-1999  
R:Accession: A44256; B44256; C44256; D44256; E44256; F44256  
R:Ragoussis, J.; Senger, G.; Mockridge, I.; Saneau, P.; Ruddy, S.; Dudley, K.; Sheer,  
Genomics 14, 673-679, 1992  
A:Title: A testis-expressed Zn finger gene (ZNF76) in human 6p21.3 centromeric to the M  
A:Reference number: A44256; MUID:93052398; PMID:1427894  
A:Accession: A44256  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-515 <RAG>  
A:Cross-references: UNIPARC:UPI000016B352; GB:M91592; NID:gl293897; PIDN:AAA98739.1; PI  
A:Experimental source: T-cell line CEM  
A:Note: sequence modified after extraction from NCBI backbone  
C:Keywords: zinc finger

Query Match 13.8%; Score 107; DB 2; Length 515;  
Best Local Similarity 24.6%; Pred. No. 0.15;  
Matches 42; Conservative 19; Mismatches 64; Indels 46; Gaps 6;

QY 9 SARQYKDLNWM-SDDKPLCTAGCGQRTN-----EDH----- 41  
DB 298 SATNKNHVRINTGKPYVCTVPGCGKRTVEYSSLYNDHVHVTCKPYTCGCKTYROT 357

QY 42 --LAVHKHKTMLKFGPARNKIEEPSVVEVTHQDPLPHE-----STTSDEKVP 92  
DB 358 STLAVHKHSAHGELEATESEQALYEQQOLEAASAESEPPKPRPRIAYLSEVKEERDDI 417

QY 93 LAQTAQTSIAIVRPASLOQVNPVLLTSSDSSVLIQAVPSPTSVTITQAPS 143

```
Db 418 PAQVAMVTE-----EDGAPQVALITQDGA-----QQVTIITSGAVVAEDSS 458
388 DSEGBESSSSIIITGGAOTPPTRLDGSAGS---SSGVSSLSGGGIKSSPHS 437

RESULT 13
T38743
hypothetical zinc finger protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T38743
R:Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, December 1995
A:Reference number: Z21808
A:Accession: T38743
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-582 <GEN>
A:Cross-references: UNIPROT:Q10076; UNIPARC:UPI000013A12B; EMBL:Z68144; PIDN:CAA92264.1;
A:Experimental source: strain 972h-; cosmid c3H1
C:Genetics:
A:Gene: SPDB:SPAC3H1.11
A:Map position: 1

Query Match 13.8%; Score 107; DB 2; Length 582;
Best Local Similarity 26.7%; Pred. No. 0.17;
Matches 39; Conservative 21; Mismatches 62; Indels 24; Gaps 6;

Qy 2 KFKLHNSARQYKDLWNMSDDKPFLLCTAPGCGQRTNEDHLAVHKHHEMTLKFGPARNS 61
||| : : : : : ||| : : : : :
Db 431 KFKRSEHLRHRISL--HTSEKPFVCI---CGKFRSRDLRQHERLH---VNASP-RLA 481
||| : : : : : ||| : : : : :
Qy 62 KTEEPSVWETHQDSPLPHEPSTTSDEKEVPLAQTAQTSIAIVRPASLOVPNVLLTSSDS 121
||| : : : : : ||| : : : : :
Db 482 CFQPSGYSSGAPGAPVQPKPIEDLNKIPINQCHD-----SSQIENLMLSSQ 532
||| : : : : : ||| : : : : :

Qy 122 SVIIQAVPSTSTVITQAPSNRP 147
: ||| : ||| : ||| : ||| : |||
Db 533 RPLSQOIVPE-----IAAYPNSIRP 552
||| : ||| : ||| : ||| : |||

RESULT 14
A49839
odd-paired - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: A49839
R:Benedyk, M.J.; Mullen, J.R.; DiNardo, S.
Genes Dev. 8, 105-117, 1994
A:Title: odd-paired: a zinc finger pair-rule protein required for the timely activation
A:Reference number: A49839; MUID:94116854; PMID:8288124
A:Accession: A49839
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-609 <BEN>
A:Cross-references: UNIPROT:P39768; UNIPARC:UPI0000130D63; GB:U04435; NID:g437303; PIDN:
C:Genetics:
A:Gene: opa
A:Cross-references: FlyBase:FBgn0003002
C:Keywords: zinc finger

Query Match 13.3%; Score 103; DB 2; Length 609;
Best Local Similarity 24.9%; Pred. No. 0.4;
Matches 43; Conservative 23; Mismatches 67; Indels 40; Gaps 7;

Qy 2 KFKLHNSARQYKDLWNMSDDKPFLLCTAPGCGQRTNEDHLAVHKHHEMTLKFGP----- 57
||| : ||| : ||| : ||| : |||
Db 275 KYKL-VNHIRVH-----TGEKPFACPHPGCGKVFARSENLIKHKRTHTGCEPKFCEHSG 327
||| : ||| : ||| : ||| : |||
Qy 58 -----ARNSKIEPSVVEVTHQ-----DSPLPHEPSTTS-----DEKEVPLAQ 95
||| : ||| : ||| : ||| : |||
Db 328 CDRFANSSDRKKHSHVHTSDPYNCRINGCDKSYTHPSLKKHKMKVGNVDEKSPSHGY 387
||| : ||| : ||| : ||| : |||
Qy 96 TAAQ----PTSAIVRPASLOVPNVLLTSSDSVIIQAVPSTSTVITQAPSS 144
||| : ||| : ||| : ||| : |||
```

Search completed: January 10, 2006, 23:06:56  
Job time : 18 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 23:00:17 ; Search time 71 Seconds  
(without alignment)  
1480.616 Million cell updates/sec

Title: AAH26175

Perfect score: 776

Sequence: 1 mkfklhvsarqykdlnms.....peptastvitqapsanrpiv 149

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05-80:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	736	94.8	209	Q8TAR1 HUMAN	Q8TAR1 homo sapien
2	736	94.8	505	Q4R5T2 MACFA	Q4R5T2 macaca fasc
3	636	82.0	234	Q8BN75 MOUSE	Q8BN75 mus musculus
4	636	82.0	420	Q8CBR9 MOUSE	Q8CBR9 mus musculus
5	636	82.0	487	1 ATF2 HUMAN	P15336 homo sapien
6	636	82.0	487	1 ATF2 MOUSE	P15951 mus musculus
7	636	82.0	487	2 O8CGB4 MOUSE	Q8CGB4 mus musculus
8	632	81.4	487	1 ATF2 RAT	Q00869 rattus norv
9	612	78.9	487	1 ATF2 CHICK	Q93602 gallus gall
10	594	76.5	486	2 Q91576 XENLA	Q91576 xenopus lae
11	445	57.3	440	2 Q68FE3 MOUSE	Q68FE3 mus musculus
12	445	57.3	447	2 Q543G2 MOUSE	Q543G2 mus musculus
13	366	47.2	389	2 Q640L6 MOUSE	Q640L6 mus musculus
14	300	38.7	413	1 ATF7 MOUSE	Q8T081 mus musculus
15	295	38.0	483	1 ATF7 PONFI	Q519C9 pongo pygma
16	289.5	37.3	413	2 Q4RXI6 TETNG	Q4RXI6 tetraodon n
17	289.5	37.3	494	1 ATF7 HUMAN	P17544 homo sapien
18	278.5	35.9	497	2 Q4VGM9 BRARE	Q4VGM9 brachydanio
19	273.5	35.2	427	2 Q4VGM8 BRARE	Q4VGM8 brachydanio
20	268.5	34.6	148	2 Q75N02 HUMAN	Q75N02 homo sapien
21	260.5	33.6	508	1 CREB5 HUMAN	Q02930 homo sapien
22	253	32.6	540	2 Q4RW6 TETNG	Q4RW6 tetraodon n
23	235	30.3	307	2 Q9UD83 HUMAN	Q9UD83 homo sapien
24	224	28.9	144	2 Q96J78 HUMAN	Q96J78 homo sapien
25	221	28.5	525	2 Q4RHY0 TETNG	Q4RHY0 tetraodon n
26	213	27.4	117	2 Q81VR8 HUMAN	Q81VR8 homo sapien
27	181.5	23.4	61	2 Q4RNR7 TETNG	Q4RNR7 tetraodon n
28	168	21.6	135	2 Q8BKN9 MOUSE	Q8BKN9 mus musculus
29	168	21.6	144	2 Q8BM42 MOUSE	Q8BM42 m mus muscu
30	159.5	20.6	571	2 Q4SIX1 TETNG	Q4SIX1 tetraodon n
31	138	17.8	23	2 Q53RY2 HUMAN	Q53RY2 homo sapien

32	135.5	17.5	226	2	Q4H3W0 CIOIN	Q4H3W0 ciona intes
33	135.5	17.5	631	2	Q4H3W1 CIOIN	Q4H3W1 ciona intes
34	116.5	15.0	396	2	Q6FMB7 CANGA	Q6FMB7 candida gla
35	116	14.9	717	2	Q55N01 CRYNE	Q55N01 cryptococcu
36	115.5	14.9	509	2	Q4PGI7 USTMA	Q4PGI7 ustilago ma
37	114.5	14.8	503	2	Q9NAR8 BRAFL	Q9NAR8 brachyosio
38	113	14.6	428	1	RSV1 SCHPO	Q9P7D9 schizosacch
39	113	14.6	568	1	Q8BMU0 MOUSE	Q8BMU0 mus musculu
40	111.5	14.4	1354	2	Q9W6B2 XENLA	Q9W6B2 xenopus lae
41	111.5	14.4	1553	2	Q6U6Z9 BRARE	Q6U6Z9 brachydanio
42	111	14.3	183	2	Q9PVD6 BRARE	Q9PVD6 brachydanio
43	111	14.3	442	2	Q93311 BRARE	Q93311 brachydanio
44	111	14.3	691	2	Q5KC36 CRYNE	Q5KC36 cryptococcu
45	110	14.2	303	2	Q6UNK6 BOVIN	Q6UNK6 bos taurus

## ALIGNMENTS

### RESULT 1

Q8TAR1 HUMAN  
ID Q8TAR1\_HUMAN PRELIMINARY; PRT; 209 AA.

AC Q8TAR1;

DT 01-JUN-2002 (Tremblrel. 21, Created)

DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)

DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)

DE ATF2 protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;

OC Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC T18SUS-Skin;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.J., Hsieh P.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Whiting M., Madan A., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield J.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

"Generation and initial analysis of more than 15,000 full-length human

and mouse cDNA sequences.";

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

FN [2]

RP NUCLEOTIDE SEQUENCE.

RC T18SUS-Skin;

RA Strausberg R.;

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

EMBL; BC026175; AAH36175.1; -, mRNA.

HSSP; P15336; 1BH1

Ensembl; ENSG00000115966; Homo sapiens.

DR GO; GO:0005634; C:nucleus; IEA.

DR GO; GO:0046872; F:metal ion binding; IEA.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR GO; GO:0008270; F:zinc ion binding; IEA.

DR InterPro; IPR007087; Znf\_C2H2.

DR Pfam; PF00096; zf-C2H2; 1.

DR SMART; SM00355; Znf\_C2H2; 1.

DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 1.

DR PROSITE; PS50157; ZINC\_FINGER\_C2H2\_2; 1.

KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.  
SQ SEQUENCE 209 AA; 23050 MW; A26AF07C45D8D5E7 CRC64;

Query Match 94.8%; Score 736; DB 2; Length 209;  
Best Local Similarity 71.3%; Pred. No. 8e-51;  
Matches 149; Conservative 0; Mismatches 0; Indels 60; Gaps 1;  
QY 1 MKFKLVNSARQYKDLWNNSDDKPFCTAPGCGQRTNEDHLAVKHGHEMTLKFGPARN 60  
DQ 1 MKFKLVNSARQYKDLWNNSDDKPFCTAPGCGQRTNEDHLAVKHGHEMTLKFGPARN 60  
QY 61 ----- 60  
DQ 61 DSVIVADQPTTFRFLKNCVEVGLFNELASPFENEFKASDDIKKMPDLDSLPLATPIIR 120  
QY 61 SKIEPSVVETTHQDSPLPHPESTTSDEKEVPLAQPTSAIVRPASLOVFNVLTTSSD 120  
DQ 121 SKIEPSVVETTHQDSPLPHPESTTSDEKEVPLAQPTSAIVRPASLOVFNVLTTSSD 180  
QY 121 SSVIIQAVPSTSSVTITQAPSSNRPIV 149  
DQ 181 SSVIIQAVPSTSSVTITQAPSSNRPIV 209

## RESULT 2

Q4RS72 MACFA PRELIMINARY; PRT; 505 AA.

AC Q4RS72; TISSUE=Eyeball;  
DT 13-SEP-2005 (TRENBLrel. 31, Created)  
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)  
DE Testis cDNA, clone: Otsa-21051, similar to human activating  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
OC Cercopitheidae; Cercopitheinae; Macaca.  
OX NCBI\_TaxID=9541;  
RN NUCLEOTIDE SEQUENCE.  
RP International consortium for macaque cDNA sequencing, analysis;  
RA "DNA sequences of macaque genes expressed in brain or testis and its  
RT evolutionary implications."  
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
RN NUCLEOTIDE SEQUENCE.  
RA Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,  
RA Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;  
RT "Substitution rate and structural divergence of 5'UTR evolution;  
RL Comparative analysis between human and cynomolgus monkey cDNAs."  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB169461; BAE01543.1; -; mRNA.  
SQ SEQUENCE 505 AA; 54507 MW; 0190EEFAC924B67 CRC64;

Query Match 94.8%; Score 736; DB 2; Length 505;  
Best Local Similarity 71.3%; Pred. No. 2.3e-50;  
Matches 149; Conservative 0; Mismatches 0; Indels 60; Gaps 1;

QY 1 MKFKLVNSARQYKDLWNNSDDKPFCTAPGCGQRTNEDHLAVKHGHEMTLKFGPARN 60  
DQ 1 MKFKLVNSARQYKDLWNNSDDKPFCTAPGCGQRTNEDHLAVKHGHEMTLKFGPARN 60  
QY 61 ----- 60  
DQ 61 DSVIVADQPTTFRFLKNCVEVGLFNELASPFENEFKASDDIKKMPDLDSLPLATPIIR 120  
QY 61 SKIEPSVVETTHQDSPLPHPESTTSDEKEVPLAQPTSAIVRPASLOVFNVLTTSSD 120  
DQ 121 SKIEPSVVETTHQDSPLPHPESTTSDEKEVPLAQPTSAIVRPASLOVFNVLTTSSD 180  
QY 121 SSVIIQAVPSTSSVTITQAPSSNRPIV 149  
DQ 181 SSVIIQAVPSTSSVTITQAPSSNRPIV 209

## RESULT 3

Q8BN75 MOUSE PRELIMINARY; PRT; 234 AA.  
AC Q8BN75; TISSUE=Eyeball;  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
DE Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched  
DE library, clone: E13020102 product: activating transcription factor 2,  
DE full insert sequence.  
GN Name=Atf2;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=C57BL/6J; TISSUE=Eyeball;  
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RX Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning."  
RL Meth. Enzymol. 303:19-44 (1999).  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=C57BL/6J; TISSUE=Eyeball;  
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RX Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RX Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,  
RX Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RX Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RX Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RX Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,  
RX Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RX Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RX Blake J., Buffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RX Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RX Guscinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RX Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RX Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RX Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RX Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RX Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RX Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690 (2001).  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=C57BL/6J; TISSUE=Eyeball;  
RC The FANTOM Consortium;  
RA The RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."  
RN Nature 420:563-573 (2002).  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=C57BL/6J; TISSUE=Eyeball;  
RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RX Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RX Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes."  
RL Genome Res. 10:1617-1630 (2000).  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=C57BL/6J; TISSUE=Eyeball;  
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RX Konno H., Akiyama J., Nishi K., Kitsumai T., Tashiro H., Itoh M.,  
RX Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,





```

RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda I., Tanaka I.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK035438; BCAC29065.1; -; mRNA.
DR HSSP; P15336; 1BHI.
DR Ensemble; ENSMUSG0000027104; Mus musculus.
DR MGI; MGI:109349; Cnt2.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR011616; bZIP_1.
DR InterPro; IPR004827; TF_bZIP.
DR InterPro; IPR007087; ZnF_C2H2.
DR Pfam; PF00096; zf-C2H2; 1.
DR SMART; SM00338; BRLZ; 1.
DR SMART; SM00355; ZnF_C2H2; 1.
DR PROSITE; PS00217; bZIP; 1.
DR PROSITE; PS00036; bZIP_BASIC; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
SQ SEQUENCE 420 AA; 45728 MW; D15I34149A971ACA CRC64;

Query Match 82.0%; Score 636; DB 2; Length 420;
Best Local Similarity 68.6%; Pred. No. 1.7e-42;
Matches 131; Conservative 0; Mismatches 0; Indels 60; Gaps 1;

QY 19 MSDDKPLCTAPCGORFTNEDHLAVHKHKTGKFGPARN----- 60
Db 1 MSDDKPLCTAPCGORFTNEDHLAVHKHKTGKFGPARNDSVIVADQTPTRFLKN 60
QY 61 -----SKIEPSVWETHQDSPL 78
Db 61 CEEVGLFNLASPFNEFKASDDIKKMPDLSPLATPIIRSKIEPSVWETHQDSPL 120
QY 79 PHPESTTSKEVPLAQTPQSAIVRPASLQVNPVLLTSSDSSVLIQAVPSPTSTVI 138
Db 121 PHPESTTSKEVPLAQTPQSAIVRPASLQVNPVLLTSSDSSVLIQAVPSPTSTVI 180
QY 139 TQAPSSNRPIV 149
Db 181 TQAPSSNRPIV 191

RESULT 5
ATP2_HUMAN
ID ATP2_HUMAN STANDARD; PRT; 487 AA.
AC P15336; Q13000;
DT 01-APR-1990 (Rel. 14, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Cyclic-AMP-dependent transcription factor ATF-2 (Activating
DE transcription factor 2) (cAMP response element binding protein CRE-
DE BP1) (HB16).
GN Name=ATF2; Synonyms=CREB2, CREBP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Fetal brain;
RX MEDLINE=90005408; PubMed=2529117;
RA Maekawa T., Sakura H., Kanei-Ishii C., Sudo T., Yoshimura T.,
RA Fujisawa J.I., Yoshida M., Ishii S.;
RT "Leucine zipper structure of the protein CRE-BP1 binding to the cyclic
RT AMP response element in brain.";
RL EMBO J. 8:2023-2028(1989).

RN NUCLEOTIDE SEQUENCE (ISOFORM 2).
RP TISSUE=Thymus;
RX MEDLINE=97211815; PubMed=9058782;
RA Yang L., Lanier B.R., Kraig E.;
RT "Identification of a novel, spliced variant of CREB that is
RT preferentially expressed in the thymus.";
RL J. Immunol. 158:2522-2525(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE OP 193-487.
RX MEDLINE=90205810; PubMed=2320002;
RA Kara C.J., Liou H.-C., Ivaehkiv L.B., Glimcher L.H.;
RT "A cDNA for a human cyclic AMP response element-binding protein which
RT is distinct from CREB and expressed preferentially in brain.";
RL Mol. Cell. Biol. 10:1347-1357(1990).
RN [4]
RP PHOSPHORYLATION SITE SER-94.
RX PubMed=15302935; DOI=10.1073/pnas.0404720101;
RA Beausoleil S.A., Jedrychowaki M., Schwartz D., Elias J.E., Villen J.,
RA Li J., Cohn M.A., Cantley L.C., Gygi S.P.;
RT "Large-scale characterization of HeLa cell nuclear phosphoproteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:12130-12135(2004).
CC -!- FUNCTION: This protein binds the cAMP response element (CRE)
CC (consensus: 5'-GTGACGTAC)[AGI-3'], a sequence present in many
CC viral and cellular promoters.
CC -!- SUBUNIT: Binds DNA as a dimer. Interacts with SMAD3 and smad4.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P15336-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P15336-2; Sequence=VSP_000587, VSP_000588;
CC -!- TISSUE SPECIFICITY: Abundant expression seen in the brain.
CC -!- PTM: Phosphorylation of Thr-51 and Thr-53 by MAPK14 causes
CC increased transcriptional activity. Also phosphorylated and
CC activated by JNK.
CC -!- SIMILARITY: Belongs to the bZIP family. ATF subfamily.
CC -!- SIMILARITY: Contains 1 bZIP domain.
CC -!- SIMILARITY: Contains 1 C2H2-type zinc finger.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X15875; CAAG33886.1; ALT_INIT; mRNA.
DR EMBL; U16028; AAB64017.1; -; mRNA.
DR EMBL; M31630; AAA35951.1; -; mRNA.
DR PIR; S05380; S05380.
DR PDB; 1BHI; NMR; 0=1-38.
DR PDB; 1T2K; X-ray; D=336-396.
DR TRANSFAC; T00167; -.
DR Ensemble; ENSG00000115966; Homo sapiens.
DR HGNC; HGNC:784; ATF2.
DR MIM; 123811; -.
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. . .; TAS.
DR GO; GO:0003713; F:transcription coactivator activity; TAS.
DR InterPro; IPR011616; bZIP_1.
DR InterPro; IPR004827; TF_bZIP.
DR InterPro; IPR007087; ZnF_C2H2.
DR Pfam; PF00170; bZIP_1; 1.
DR Pfam; PF00096; zf-C2H2; 1.
DR SMART; SM00338; BRLZ; 1.
DR SMART; SM00355; ZnF_C2H2; 1.
DR PROSITE; PS00217; bZIP; 1.
DR PROSITE; PS00036; bZIP_BASIC; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
KW 3D-structure; Activator; Alternative splicing; DNA-binding;
KW Metal-binding; Nuclear protein; Phosphorylation; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.

```

```

FT DOMAIN 362 390 Leucine-zipper.
FT ZN_FING 7 31 C2H2-type.
FT DN_BIND 33 356 Basic motif.
FT MOD_RES 51 51 Phosphothreonine (by MAPK14).
FT MOD_RES 53 53 Phosphothreonine (by MAPK14).
FT MOD_RES 94 94 Phosphoserine.
FT VARSPLIC 1 158 Missing (in isoform 2).
FT VARSPLIC /FTID=VSP_000587.
FT VARSPLIC 159 167 TSSDSSVII -> MSTAYFQM (in isoform 2).
FT CONFLICT 191 191 V -> L (in Ref. 2).
FT CONFLICT 205 205 N -> S (in Ref. 1).
FT CONFLICT 293 293 R -> L (in Ref. 2).
FT TURN 2 3
FT STRAND 7 8
FT TURN 12 14
FT STRAND 17 18
FT HELIX 21 32
FT TURN 33 37
SQ SEQUENCE 487 AA; 52277 MW; 58ADD6240D6270E8 CRC64;

Query Match 82.0%; Score 636; DB 1; Length 487;
Best Local Similarity 68.6%; Pred. No. 2e-42;
Matches 131; Conservative 0; Mismatches 0; Indels 60; Gaps 1;

QY 19 MSDDKPFLLCTAGCGQRTNEDHLAVHKHHEMTLKFQPARN----- 60
DB 1 MSDDKPFLLCTAGCGQRTNEDHLAVHKHHEMTLKFQPARNDSVIVADQTPTRFLKN 60
QY 61 -----SKTEPSVVEVTHQDSPL 78
DB 61 CEEVGLFNLASPFENEFKASEDDIKKPLDLSPLATPIIRSKIEPSVVEVTHQDSPL 120
QY 79 PHPESTTSDEKEVPLAQTAOPTSAIVRPASLOVPNVLLTSSDSSVIIQOAVPSPSTTVI 138
DB 121 PHPESTTSDEKEVPLAQTAOPTSAIVRPASLOVPNVLLTSSDSSVIIQOAVPSPSTTVI 180
QY 139 TQAPSSNRPIV 149
DB 181 TQAPSSNRPIV 191

RESULT 6
ATF2_MOUSE STANDARD; PRT; 487 AA.
AC P16951; Q64089; Q64090; Q64091;
DT 01-AUG-1990 (Rel. 15, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Cyclic-AMP-dependent transcription factor ATF-2 (Activating
DE transcription factor 2) (cAMP response element binding protein CRE-
DE BPI) (NBXP protein).
GN Name=Atf2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OK NCBI_TaxID=10090;
[1]
RN RP NUCLEOTIDE SEQUENCE OF 9-487, AND ALTERNATIVE SPLICING.
RX MEDLINE=92123199; PubMed=1531087;
RA Georgopoulos K., Morgan B.A., Moore D.D.;
RT "Functionally distinct isoforms of the CRE-BP DNA-binding protein
RL mediate activity of a T-cell-specific enhancer.";
RL Mol. Cell. Biol. 12:747-757(1992).
[2]
RN RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC STRAIN=ILS, and ISS;
RX MEDLINE=21363810; PubMed=11471062; DOI=10.1007/s00335-001-1001-x;
RA Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Cammiff J.,
RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikela J.M.;
RT "High-throughput sequence identification of gene coding variants
RT within alcohol-related QTLs.";

```

```

Mamm. Genome 12:657-663 (2001).
[3]
RN RP NUCLEOTIDE SEQUENCE OF 77-487.
RX MEDLINE=90205841; PubMed=2138707;
RA Ivashkiv L.B., Liou H.-C., Kara C.J., Lamph W.W., Verma I.M.,
RA Glimcher L.H.;
RT "MXBP/CRE-BP2 and c-Jun form a complex which binds to the cyclic AMP,
RT but not to the 12-O-tetradecanoylphorbol-13-acetate, response
RT element.";
RL Mol. Cell. Biol. 10:1609-1621(1990).
CC -!- FUNCTION: A transcriptional activator, probably constitutive,
CC which binds to the cAMP responsive region (CRE) (consensus: 5'-
CC GTGACGT[AC][AG]-3'); a sequence present in many viral and cellular
CC promoters. The interaction of MXBP/CRE-BP2 with JUN redirecs JUN
CC to bind to CREs preferentially over the 12-O-tetradecanoylphorbol-
CC 13-acetate response elements (TREs) as part of an MXBP-c-Jun
CC complex.
CC -!- SUBUNIT: Binds DNA as a dimer and can form a homodimer in the
CC absence of DNA. Can form an heterodimer with JUN.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=P16951-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P16951-2; Sequence=VSP_000590;
CC Name=3;
CC IsoId=P16951-3; Sequence=VSP_000589;
CC -!- SIMILARITY: Belongs to the bZIP family. ATF subfamily.
CC -!- SIMILARITY: Contains 1 bZIP domain.
CC -!- SIMILARITY: Contains 1 C2H2-type zinc finger.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
EMBL; S76657; AAB21128.1; ALT_INIT; mRNA.
DR EMBL; S76659; AAB21129.1; ALT_INIT; mRNA.
DR EMBL; S76655; AAB21127.1; -; mRNA.
DR EMBL; AF483482; AAL90756.1; -; mRNA.
DR EMBL; AF483483; AAL90757.1; -; mRNA.
DR EMBL; M31629; AAA39780.1; -; mRNA.
DR HSSP; P15336; 1BHI.
DR SMR; P16951; 336-396.
DR TRANSFAC; T01017; -.
DR Ensembl; ENSMUSG00000027104; Mus musculus.
DR MGI; MGI:109349; Atf2.
GO; GO:0003677; F:DNA binding; IDA.
GO; GO:0005515; F:protein binding; IPI.
InterPro; IPR011616; bZIP_1.
InterPro; IPR004827; TF_bZIP.
InterPro; IPR007087; ZnF_C2H2.
Pfam; PF00170; bZIP_1; 1.
Pfam; PF00096; zf-C2H2; 1.
SMART; SM00338; BRL2; 1.
SMART; SM00355; ZnF_C2H2; 1.
PROSITE; PS0217; bZIP_1.
PROSITE; PS00036; bZIP_BASIC; 1.
PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
KW Activator; Alternative splicing; DNA-binding; Metal-binding;
KW Nuclear protein; Phosphorylation; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
FT DOMAIN 362 390 Leucine-zipper.
FT ZN_FING 7 31 C2H2-type.
FT DN_BIND 33 356 Basic motif.
FT MOD_RES 51 51 Phosphothreonine (by MAPK14) (By
FT similarity).
FT MOD_RES 53 53 Phosphothreonine (by MAPK14) (By
FT similarity).
FT MOD_RES 94 94 Phosphoserine (By similarity).

```

FT	VARSPLIC	1	48
FT	MSDDKPFCLTAPGCCGQRTNEDHLAVHKKHEMTLKFGPAR		NDSVIVA -> MHCPTWMP (in isoform 3).
FT	/FTID=VSP_000589.		
FT	VARSPPLIC	132	229
FT	Missing (in isoform 2).		
FT	/FTID=VSP_000590.		
FT	CONFLICT	482	487
FT	AOPSGS -> HSPOEVD (in Ref. 1).		
SQ	SEQUENCE	487 AA;	52298 MW; F9CDEC3BC3119ACB CRC64;
 Query Match Best Local Similarity 68.6%; DB 1; Length 487; Matches 131; Conservative 0; Mismatches 0; Indels 60; Gaps 1;			
Qy	19 MSDDKPFCLTAPGCCGQRTNEDHLAVHKKHEMTLKFGPAR-		----- 60
Dd	1 MSDDKPFCLTAPGCCGQRTNEDHLAVHKKHEMTLKFGPARNDSVIVADQTPTTRFLKN		60
Qy	61 -----		-----SKIEBPVVETTHODSPL 78
Dd	61 CEEVGLFNELASPENEFFKAASEDDIKOMPLDLSPLATPIIRSKIEBPVVETTHODSPL		120
Qy	79 PHESTTTSDKEVEPLAQTAOPTSAIVRPASLQVPNVLLTSSDSSVIIQQAVPSPTSSTVI		138
Dd	121 PHESTTTSDKEVEPLAQTAOPTSAIVRPASLQVPNVLLTSSDSSVIIQQAVPSPTSSTVI		180
Qy	139 TQAPSSNRPIV 149		
Dd	181 TQAPSSNRPIV 191		
 RESULT 7			
ID	Q8CGB4 MOUSE PRELIMINARY;	PRT;	487 AA.
OC	Q8CGB4; MOUSE		
AC	01-MAR-2003 (TrEMBLrel. 23, Created)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)		
DE	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)		
DE	Atf2 protein.		
GN	Name=Atf2;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		
OC	Muridae; Murinae; Mus.		
OX	NCBI_Taxid=10090;		
RN	[1]		
RC	NUCLEOTIDE SEQUENCE.		
RP	STRAIN=CZECH II;		
RC	TISSUE=Mammary tumor metastatized to lung. Tumor arose spontaneously;		
RX	MDLINE=2238825; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Udén T.B., Toshitsuki S., Carninci P., Frange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalón D.K., Muzy D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,		
RA	Faney J., Helton E., Kettman M., Madan A.C., Shevchenko Y., Bouffard G.G.,		
RA	Whiting M., Madan A., Young A.C., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Touchman J.W., Schmutz J., Myers R.M.,		
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,		
RA	Schnierch A., Schein J.E., Jones S.J.M., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RL	tissue cDNA sequences."		
RU	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=CZECH II;		
RC	TISSUE=Mammary tumor metastatized to lung. Tumor arose spontaneously;		
RA	Straussberg R.		

RL	Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR	EMBL; BC042210; AH42210.1; -, mRNA.
DR	HSSP; P15336; IBHI.
DR	SNR; Q8CGB4; 336-396.
DR	MGI; MGI:109349; Atf2.
DR	GO; GO:0003677; F:DNA binding; IDA.
DR	GO; GO:0005515; F:protein binding; IPI.
DR	InterPro; IPRO11615; BZIP.1.
DR	InterPro; IPRO04827; TF_BZIP.
DR	InterPro; IPRO07087; ZnF_C2H2.
DR	Pfam; PF00170; bZIP.1; 1.
DR	Pfam; PF00096; zf-C2H2; 1.
DR	SMART; SM00338; BRLZ; 1.
DR	SMART; SM00355; ZnF_C2H2; 1.
DR	PROSITE; PS00217; BZIP.1.
DR	PROSITE; PS00036; BZIP_BASIC; 1.
DR	PROSITE; PS00028; ZINC_FINGER_C2H2.1; 1.
DR	PROSITE; PS00157; ZINC_FINGER_C2H2.2; 1.
DR	SEQUENCE 487 AA; 5310 MW; FA2B093BC210CFCEB CRC64;
QY	Query Match 82.0%; Score 636; DB 2; Length 487;
DB	Best Local Similarity 68.6%; Pred. No. 2e-42;
DB	Matches 131; Conservative 0; Mismatches 0; Indels 60; Gaps 1
QY	19 MSDDKDFLCTAPCGGQFTNEDHLAVHKHKHEMTLKFGPARN----- 60
DB	1 MSDDKDFLCTAPCGGQFTNEDHLAVHKHKHEMTLKFGPARN----- 60
QY	61 -----SKIEEPSVVETTHQDSPL 78
DB	61 CEEVGLFNLASPFENEFKASEDDICKMPLDLSPLATPIIRSKIEEPSVVETTHQDSPL 120
QY	79 PHESTTSDSEKVEPLAQTAQTSATVIRPASLOQPNVLLTSSDSVVIQQAVPSPTSSTVI 138
DB	121 PHESTTSDSEKVEPLAQTAQTSATVIRPASLOQPNVLLTSSDSVVIQQAVPSPTSSTVI 180
QY	139 TQAPSSNRPIV 149
DB	181 TQAPSSNRPIV 191
RESULT 8	
ID	ATF2_RAT STANDARD; PRT; 487 AA.
AC	Q00969; Q62870;
DT	01-APR-1993 (Rel. 25, Created)
DT	30-MAY-2000 (Rel. 39, Last sequence update)
DT	13-SEP-2005 (Rel. 48, Last annotation update)
DE	Cyclic-AMP-dependent transcription factor ATF-2 (Activating
DE	transcription factor 2) (cAMP response element binding protein CRE-
DE	Bp1).
GN	Name=Atf2;
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC	Muroidea; Muridae; Murinae; Rattus.
OX	NCBI_TaxID=10116;
RN	[1]
RP	NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC	STRAIN=Sprague-Dawley; TISSUE=Brain;
RA	Muramatsu S.;
RL	Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN	[2]
RP	NUCLEOTIDE SEQUENCE (ISOFORM 2).
RC	TISSUE=Brain;
RX	MEDLINE=91332085; PubMed=1714459;
RT	Kageyama R., Sasai Y., Nakanishi S.;
RT	"Molecular characterization of transcription factors that bind to the
RT	cAMP responsive region of the substance P precursor gene. cDNA cloning
RT	of a novel C/EBP-related factor."
RT	J. Biol. Chem. 266:152525-15531(1991).
CC	-1- FUNCTION: transcriptional activator, probably constitutive, which
CC	binds to the cAMP response element (CRE) (consensus: 5'-

```
CC GTGACGT[AC][AG]-3'), a sequence present in many viral and cellular
CC promoters.
CC -|- SUBUNIT: Binds DNA as a dimer.
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q00969-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q00969-2; Sequence=VSP_000591;
CC -|- SIMILARITY: Belongs to the bZIP family. ATF subfamily.
CC -|- SIMILARITY: Contains 1 bZIP domain.
CC -|- SIMILARITY: Contains 1 C2H2-type zinc finger.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U38938; AA93263.1; -; mRNA.
CC EMBL; M65148; AAA42013.1; -; mRNA.
CC PIR; A39429; A39429.
CC HSSP; P15336; 1BH1.
CC SMR; O00969; 336-396.
CC TRANSFAC; T01382; -.
CC ENSEMBL; ENSRNOG0000001597; Rattus norvegicus.
CC RGD; 621862; Atf2.
CC GO; GO:0016563; P:transcriptional activator activity; IDA.
CC GO; GO:0042423; P:catecholamine biosynthesis; IPI.
CC GO; GO:0007399; P:neurogenesis; NAS.
CC InterPro; IPR011616; bZIP_1.
CC InterPro; IPR004827; TF_bZIP.
CC InterPro; IPR007087; ZnF_C2H2.
CC Pfam; PF00170; bZIP_1; 1.
CC Pfam; PF00096; zf-C2H2; 1.
CC PROSITE; PS0217; bZIP_1.
CC PROSITE; PS00036; bZIP_BASIC; 1.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
CC Activator; Alternative splicing; DNA-binding; Metal-binding;
CC Nuclear protein; Phosphorylation; Transcription;
CC Transcription regulation; Zinc; Zinc-finger.
CC DOMAIN 362 390 Leucine-zipper.
CC ZN_FING 7 31 C2H2-type.
CC MOD_RES 53 53 Phosphothreonine (by MAPK14) (By
FT MOD_RES 94 94 Phosphoserine (By similarity).
FT MOD_RES 132 229 Missing (in isoform 2).
FT FTID=VSP_000591.
CC SEQUENCE 487 AA; 52287 MW; 4ED95B106DF5F9EE CRC64;
Query Match 81.4%; Score 632; DB 1; Length 487;
Best Local Similarity 67.5%; Pred. No. 4,1e-42;
Matches 129; Conservative 2; Mismatches 0; Indels 60; Gaps 1;
QY 19 MSDDKPFLLCTAPGCGQRTNEDHLAVHKHHEMTLKFGPARN----- 60
DB 1 MSDDKPFLLCTAPGCGQRTNEDHLAVHKHHEMTLKFGPARNDSVIVADQTPPTRFLEKN 60
QY 61 -----SKIEPSVVETTHQDSPL 78
DB 61 CEEVGLFNLASPFENEFKASDDIKKPLDLSPLATPIIRSKIEPSVVETTHQDSPL 120
QY 79 PHPESTTSDEKEVPLAQTAQPTSAIVRPASIQVNPVLLTSSDSSVLIQQAVPSPSTSTVI 138
DB 121 PHPESTTNDEKEIPLAQTAQPTSAIVRPASIQVNPVLLTSSDSSVLIQQAVPSPSTSTVI 180
QY 139 TOAPSSNRPIV 149
|||||
```

```
Db 181 TOAPSSNRPIV 191
RESULT 9
ATF2_CHICK
ID ATF2_CHICK STANDARD; PRT; 487 AA.
AC O93602;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Cyclic-AMP-dependent transcription factor ATF-2 (Activating
DE transcription factor 2).
GN Name=ATF2;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=95038207; PubMed=9819389;
RA Huguier S., Baguet J., Perez S., van Dam H., Castellazzi M.;
RT "Transcription factor ATF2 cooperates with v-Jun to promote growth
RT factor-independent proliferation in vitro and tumor formation in
RT vivo.";
RL Mol. Cell. Biol. 18:7020-7029(1998).
CC -|- FUNCTION: Transcriptional activator, probably constitutive, which
CC binds to the cAMP response element (CRE) (consensus: 5'-
CC GTGACGT[AC][AG]-3'), a sequence present in many viral and cellular
CC promoters (By similarity).
CC -|- SUBUNIT: Binds DNA as a dimer.
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- SIMILARITY: Belongs to the bZIP family. ATF subfamily.
CC -|- SIMILARITY: Contains 1 bZIP domain.
CC -|- SIMILARITY: Contains 1 C2H2-type zinc finger.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; Y17724; CAA76838.1; -; mRNA.
CC HSSP; P15336; 1BH1.
CC SMR; O93602; 336-396.
CC ENSEMBL; ENSGALG00000009287; Gallus gallus.
CC InterPro; IPR011616; bZIP_1.
CC InterPro; IPR004827; TF_bZIP.
CC InterPro; IPR007087; ZnF_C2H2.
CC Pfam; PF00170; bZIP_1; 1.
CC Pfam; PF00096; zf-C2H2; 1.
CC SMART; SM00338; BRZ; 1.
CC SMART; SM00355; ZnF_C2H2; 1.
CC PROSITE; PS0217; bZIP_1.
CC PROSITE; PS00036; bZIP_BASIC; 1.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
CC Activator; DNA-binding; Metal-binding; Nuclear protein; Transcription;
CC Transcription regulation; Zinc; Zinc-finger.
CC DOMAIN 362 390 Leucine-zipper.
CC ZN_FING 7 31 C2H2-type.
CC MOD_RES 333 356 Basic motif.
CC SEQUENCE 487 AA; 52407 MW; A1F42734D9C6A146 CRC64;
Query Match 78.9%; Score 612; DB 1; Length 487;
Best Local Similarity 65.4%; Pred. No. 1.6e-40;
Matches 125; Conservative 3; Mismatches 3; Indels 60; Gaps 1;
QY 19 MSDDKPFLLCTAPGCGQRTNEDHLAVHKHHEMTLKFGPARN----- 60
DB 1 MSDDKPFLLCTAPGCGQRTNEDHLAVHKHHEMTLKFGPARNDSVIVADQTPPTRFLEKN 60
QY 61 -----SKIEPSVVETTHQDSPL 78
```

```
Db 61 CEEVGLFNLASPFENEFKASDDIKQMPDLSPATPIIRNKIEPSVETTHQDSPL 120
QY 79 PHESTTSDEKVPPLAQTAQTSATVRPASLQVNPVLLTSSDSSVIIQQAIVPSPTSTVI 138
Db 121 PHESTTNDKEVSLQTAQTSATVRPASLQVNPVLLTSSDSSVIIQQAIVPSPTSTVI 180
QY 139 TQAPSSNRPIV 149
Db 181 TQAPSSNRPIV 191
RESULT 10
Q91576_XENLA
ID Q91576_XENLA PRELIMINARY; PRT; 486 AA.
AC Q91576;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Activating transcription factor 2.
GN Name=Atf2;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=95180723; PubMed=7875593; DOI=10.1016/0378-1119(94)00770-S;
RA Villarreal X.C., Richter J.D.;
RT "Analysis of Arf2 gene expression during early Xenopus laevis
RT development."
RL Gene 153:225-229(1995).
DR EMBL; U16158; AAA69518.1; -; mRNA.
DR PIR; JC4028; JC4028.
DR HSSP; P15336; 1BHI.
DR SMR; Q91576; 336-396.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR011616; bZIP 1.
DR InterPro; IPR002112; LeuZip Jun.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP 1; 1.
DR Pfam; PF00096; zf-C2H2; 1.
DR PRINTS; PR00043; LEUZIPRJUN.
DR SMART; SM00338; BRLZ; 1.
DR SMART; SM00355; ZNF_C2H2; 1.
DR PROSITE; PS50217; bZIP; 1.
DR PROSITE; PS00036; bZIP_BASIC; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
SQ SEQUENCE 486 AA; 53378 MW; 26C273C0917D17DD CRC64;
Query Match 76.5%; Score 594; DB 2; Length 486;
Best Local Similarity 64.9%; Pred. No. 4,4e-39;
Matches 124; Conservative 4; Mismatches 3; Indels 60; Gaps 2;
QY 19 MSDDKPLCFLTCPCGGORFTNEDHLAVHKHKHMTLKFPGARN----- 60
Db 1 MSDDKPLCFLTCPCGGORFTNEDHLAVHKHKHMTLKFPGARNDSIVADOTPTPTFLKN 60
QY 61 -----SKIEPSVETTHQDSPLP 79
Db 61 CEEVGLFNLASPFENEFKASDDIKQMPDLSPATPIIRNKIEPSVETTHQDSPLP 120
QY 80 HPESTTSDE-KVPLAQTAQTSATVRPASLQVNPVLLTSSDSSVIIQQAIVPSPTSTVI 138
Db 121 HPESTTSDQHNQVPLAQTAQTSATVRPASLQVNPVLLTSSDSSVIIQQAIVPSPTSTVI 180
QY 139 TQAPSSNRPIV 149
```

```
Db 181 TQAPSTNSSIV 191
RESULT 11
Q68FE3_MOUSE
ID Q68FE3_MOUSE PRELIMINARY; PRT; 440 AA.
AC Q68FE3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Atf2 protein.
GN Name=Atf2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388957; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC079883; AAH79883.1; -; mRNA.
DR SMR; Q68FE3; 289-349.
DR MGI; MGI:109349; Atf2.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR011616; bZIP 1.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP 1; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS50217; bZIP; 1.
DR PROSITE; PS00036; bZIP_BASIC; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 440 AA; 47237 MW; 65FB5785BA82B963 CRC64;
Query Match 57.3%; Score 445; DB 2; Length 440;
Best Local Similarity 70.1%; Pred. No. 2.8e-27;
Matches 103; Conservative 1; Mismatches 29; Indels 14; Gaps 4;
QY 17 WMSDDPK-----FL--CTAPG----CGQFTNEDHLAVHKHKHMTLKFPGAR----NSK 62
Db 5 WWPQDQTPPTPTFLKNCEEVGLFNLASPFENEFKASDDIKQMPDLSPATPIIRSK 64
QY 63 IEPSVETTHQDSPLPHESTTSDEKVPPLAQTAQTSATVRPASLQVNPVLLTSSDSS 122
Db 65 IEPSVETTHQDSPLPHESTTSDEKVPPLAQTAQTSATVRPASLQVNPVLLTSSDSS 124
```

QY 123 VIIQAVPSPTSTVITQAPSSNRPIV 149  
 DB 125 VIIQAVPSPTSTVITQAPSSNRPIV 151

RESULT 12  
 Q543G2 MOUSE  
 ID Q543G2 MOUSE PRELIMINARY; PRT; 447 AA.  
 AC Q543G2  
 DT 13-SEP-2005 (TREMBlrel. 31, Created)  
 DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)  
 DE 12 days embryo spinal ganglion cDNA, RIKEN full-length enriched  
 DE library, clone:D130059A07 product:activating transcription factor 2,  
 full insert sequence.  
 GN Name=Atf2;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Caeavant T.,  
 RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., King B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RL "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 403:685-690(2001).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaide I., Ohtsuka N., Saito R., Suzuki H., Yamanaoka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grammond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed J.C., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shinada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,

Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wellie C.,  
 Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RT Nature 420:563-573(2002).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RT Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwara S., Inoue K., Izawa M., Ohara E., Watanabe M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multipipillary sequencer.";  
 RT Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK051591; BAC34684.1; -; mRNA.  
 DR MGI; MGI:109349; Atf2.  
 DR GO; GO:0003677; F:DNA binding; IDA.  
 DR GO; GO:0005515; F:protein binding; IPI.  
 KW DNA-binding; Nuclear protein.  
 SQ SEQUENCE 447 AA; 48002 MW; 35A0751C1627D682 CRC64;

Query Match 57.3%; Score 445; DB 2; Length 447;  
 Best Local Similarity 70.1%; Pred. No. 2.9e-27;  
 Matches 103; Conservative 1; Mismatches 23; Indels 14; Gaps 4;  
 QY 17 WMSDDKP---FL--CTAPG---CGQFTNEDHLAVHKHKHMTLKFQPAR----NSK 62  
 DB 5 WVPDQPTPTPTFLKNCNEVGLFNELASPFNEPKKASEDDIKKMPDLSPALPIIRSK 64  
 QY 63 IEESVVTTHQDSPLPHPESTTSDEKVPVLAQTAQPTSAIVRPASLPVNPVLLTSSDSS 122  
 DB 65 IEESVVTTHQDSPLPHPESTTSDEKVPVLAQTAQPTSAIVRPASLPVNPVLLTSSDSS 124  
 QY 123 VIIQAVPSPTSTVITQAPSSNRPIV 149  
 DB 125 VIIQAVPSPTSTVITQAPSSNRPIV 151



```

Query Match      47.2%; Score 366; DB 2; Length 389;
Best Local Similarity 45.5%; Pred. No. 4.8e-21;
Matches 80; Conservative 6; Mismatches 4; Indels 86; Gaps 4
QY      19 MSDDKPFCLTAPGGCGQRTFNEDHLAVHKHKHEMTLKFGPARN----- 60
      |||||
Db       1 MSDDKPFCLTAPGGCGQRTFNEDHLAVHKHKHEMTLKFGPARNDSVIVADQTPTRFLKN 60
      |||||
QY      61 -----SKIEEPSVYVETTHQDSPL 78
      |||||

```



```
DR PROSITE; PS00170; BZIP; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 1.
DR SMART; SM00338; BRL2; 1.
DR SMART; SM00355; ZNF_C2H2; 1.
DR PROSITE; PS00217; BZIP; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 1.
KW Activator; DNA-binding; Metal-binding; Nuclear protein;
KW Phosphorylation; Transcription; Transcription regulation; Zinc;
KW Zinc-finger.
FT DOMAIN 332 395 Leucine-zipper.
FT ZN_FING 7 31 C2H2-type.
FT COMBIAS 118 253 Pro-rich.
FT COMBIAS 322 325 Poly-Arg.
SQ SEQUENCE 413 AA; 44608 MW; AFF4D1A7BFE71AF3 CRC64;

Query Match 38.7%; Score 300; DB 1; Length 413;
Best Local Similarity 38.7%; Pred. No. 9.1e-16;
Matches 76; Conservative 11; Mismatches 33; Indels 76; Gaps 6;

QY 19 MSDDKPFCTAGCGQRTNEDHLAVHKHMTLKFGPARNSKI----- 63
DB 1 MGDDRPFVCSAPCGQRTNEDHLAVHKHMTLKFGPARTDSVIIADQTPTRFLKN 60
QY 64 -----EEPSVETTHOD 75
DB 61 CEEVGLFNELASFEHFKASDDDEKGAAGPLDMSLPSTPDIKIKEEPVEVDSSPPD 120
QY 76 SPLPHPESTTSDEKVEV---PLAQTAQPTSAIVRPASLQVNPVLLTSSDSSVIIQOAVPSP 132
DB 121 SPASSPCSPPLKEKEVTPKPVV-ISTPTTIVRPGSLP-----LHLGYDP---LHPTLPSP 172
QY 133 TSSTVITQAPSSNRPI 148
DB 173 TS--VITQAPPSNRQI 186

RESULT 15
ATF7_PONFY
ID ATF7_PONFY STANDARD; PRT; 483 AA.
AC QSR9C9;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DE 13-SEP-2005 (Rel. 48, Last annotation update)
DE Cyclic-AMP-dependent transcription factor ATF-7 (Activating
DE transcription factor 7) (Transcription factor ATF-A).
GN Name=ATF7;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1] _TaxID=9600;
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Heart;
RG The German cDNA consortium;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Binds the cAMP response element (CRE) (consensus: 5'-
CC GTGACGT[AG][AG]-3'), a sequence present in many viral and cellular
CC promoters (By similarity).
CC -!- SUBUNIT: Binds DNA as a dimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the bZIP family.
CC -!- SIMILARITY: Contains 1 bZIP domain.
CC -!- SIMILARITY: Contains 1 C2H2-type zinc finger.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; CR859460; CAH91631.1; -; mRNA.
DR SMR; QSR9C9; 334-394.
DR InterPro; IPR011616; bZIP 1.
DR InterPro; IPR004827; TF_bZIP.
```

```
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00170; bZIP 1; 1.
DR Pfam; PF00096; zf-C2H2; 1.
DR SMART; SM00338; BRL2; 1.
DR SMART; SM00355; ZNF_C2H2; 1.
DR PROSITE; PS00217; BZIP; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 1.
KW Activator; DNA-binding; Metal-binding; Nuclear protein;
KW Phosphorylation; Transcription; Transcription regulation; Zinc;
KW Zinc-finger.
FT DOMAIN 332 395 Leucine-zipper.
FT ZN_FING 7 31 C2H2-type.
FT COMBIAS 118 253 Pro-rich.
FT COMBIAS 322 325 Poly-Arg.
SQ SEQUENCE 483 AA; 51743 MW; D05D084FC73330D6 CRC64;

Query Match 38.0%; Score 295; DB 1; Length 483;
Best Local Similarity 38.7%; Pred. No. 2.8e-15;
Matches 75; Conservative 10; Mismatches 33; Indels 76; Gaps 6;

QY 19 MSDDKPFCTAGCGQRTNEDHLAVHKHMTLKFGPARNSKI----- 63
DB 1 MGDDRPFVCSAPCGQRTNEDHLAVHKHMTLKFGPARTDSVIIADQTPTRFLKN 60
QY 64 -----EEPSVETTHOD 75
DB 61 CEEVGLFNELASFEHFKAADEKKAAGPLDMSLPSTPDIKIKEEPVEVDSSPPD 120
QY 76 SPLPHPESTTSDEKVEV---PLAQTAQPTSAIVRPASLQVNPVLLTSSDSSVIIQOAVPSP 132
DB 121 SPASSPCSPPLKEKEVTPKPVV-ISTPTTIVRPGSLP-----LHLGYDP---LHPTLPSP 172
QY 133 TSSTVITQAPSSNR 146
DB 173 TS--VITQAPPSNR 184

Search completed: January 10, 2006, 23:06:30
Job time : 73 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 23:02:23 ; Search time 23 Seconds

(without alignments)  
535.595 Million cell updates/sec

Title: Aah26175

Perfect score: 776

Sequence: 1 mkfklhvnarsqykdlwnms.....pepsstvtgapssnrpiv 149

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/iaa/5 COMB.pep:\*

2: /cgn2\_6/ptodata/1/iaa/6 COMB.pep:\*

3: /cgn2\_6/ptodata/1/iaa/H COMB.pep:\*

4: /cgn2\_6/ptodata/1/iaa/PCRTUS COMB.pep:\*

5: /cgn2\_6/ptodata/1/iaa/RE COMB.pep:\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	736	94.8	501	2	US-09-949-016-11281
2	212	27.3	52	2	US-09-513-999C-7023
3	137	17.7	29	2	US-09-503-954A-20
4	137	17.7	29	2	US-09-970-515A-20
5	115	14.8	27	2	US-08-819-177-6
6	109	14.0	447	2	US-09-234-332-11
7	106.5	13.7	97	2	US-09-395-448-15
8	106.5	13.7	97	2	US-09-925-796-15
9	106.5	13.7	97	2	US-09-941-450-15
10	103	13.3	522	2	US-09-949-016-11145
11	102.5	13.2	496	1	US-08-224-482-2
12	102.5	13.2	533	1	US-08-040-548-1
13	102.5	13.2	533	1	US-08-466-344-1
14	102.5	13.2	533	2	US-09-546-013-21
15	102.5	13.2	533	6	5206152-2
16	102.5	13.2	1367	2	US-09-487-558B-108
17	102	13.1	409	2	US-09-029-755C-2
18	102	13.1	443	2	US-09-342-325C-42
19	102	13.1	443	2	US-10-244-367-42
20	100	12.9	383	2	US-09-248-796A-19589
21	98.5	12.7	380	2	US-09-248-796A-19662
22	97.5	12.6	94	2	US-09-229-007A-9
23	97.5	12.6	94	2	US-10-113-424-9
24	97.5	12.6	94	2	US-09-716-637-16
25	97.5	12.6	168	1	US-08-353-476-73
26	97.5	12.6	176	2	US-09-248-796A-18892
27	97.5	12.6	241	1	US-08-353-476-116

28	97.5	12.6	273	1	US-08-353-476-112	Sequence 112, App
29	97.5	12.6	785	2	US-09-538-092-872	Sequence 872, App
30	97	12.5	99	2	US-09-229-037-17	Sequence 17, Appl
31	97	12.5	99	2	US-08-478-681-17	Sequence 17, Appl
32	97	12.5	99	2	US-09-779-233-5	Sequence 5, Appl
33	97	12.5	99	2	US-09-706-243A-17	Sequence 17, Appl
34	97	12.5	99	2	US-09-942-087A-17	Sequence 17, Appl
35	97	12.5	196	2	US-09-229-037-30	Sequence 30, Appl
36	97	12.5	196	2	US-08-478-681-30	Sequence 30, Appl
37	97	12.5	196	2	US-09-779-233-18	Sequence 18, Appl
38	97	12.5	196	2	US-09-706-243A-30	Sequence 30, Appl
39	97	12.5	196	2	US-09-942-087A-30	Sequence 30, Appl
40	95.5	12.3	441	2	US-09-172-045-2	Sequence 2, Appl
41	95.5	12.3	441	2	US-09-342-325C-2	Sequence 2, Appl
42	95.5	12.3	441	2	US-10-244-367-2	Sequence 8, Appl
43	95	12.2	496	2	US-10-164-595-8	Sequence 60, Appl
44	95	12.2	496	2	US-10-164-595-8	Sequence 44, Appl
45	95	12.2	501	2	US-09-342-325C-44	

#### ALIGNMENTS

#### RESULT 1

US-09-949-016-11281  
; Sequence 11281, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11281  
; LENGTH: 501  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11281

Query Match		94.8%	Score 736;	DB 2;	Length 501;
Best Local Similarity		71.3%	Pred. No. 4.3e-71;		
Matches 149;		Conservative 0;	Mismatches 0;	Indels 60;	Gaps 1;
QY	1	MKFKLHVNARSQYKDLWNMSDDKPFCLTAPGCGGRTNEDHLAVHKHKHMTLKFGPARN	60		
Db	5	MKFKLHVNARSQYKDLWNMSDDKPFCLTAPGCGGRTNEDHLAVHKHKHMTLKFGPARN	64		
QY	61	-----	60		
Db	65	DSVIVADQTPPTFLKNCVEGVLNFELASPFNEFKKASEDDIKKMFLLDLSPLATPIIR	124		
QY	61	SKIEEPSVVEVTHQDSPLPHEPSTTSDEKVEPLAQTAQTSIAIVRPASLQVNVLLTSSD	120		
Db	125	SKIEEPSVVEVTHQDSPLPHEPSTTSDEKVEPLAQTAQTSIAIVRPASLQVNVLLTSSD	184		
QY	121	SSVITQQAQVPSPTSTVITQAPSSNRPIV	149		
Db	185	SSVITQQAQVPSPTSTVITQAPSSNRPIV	213		

#### RESULT 2

US-09-513-999C-7023  
; Sequence 7023, Application US/09513999C  
; Patent No. 6783961

GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Duclert, A.Y.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
Patent No. 6783961  
FILE REFERENCE: 59.US2.REG  
CURRENT APPLICATION NUMBER: US/09/513,999C  
CURRENT FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/122,487  
PRIOR FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 36681  
SOFTWARE: Patent.pm  
SEQ ID NO 7023  
LENGTH: 52  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: 6  
OTHER INFORMATION: Xaa=Pro or Ser  
US-09-513-999C-7023

Query Match 27.3%; Score 212; DB 2; Length 52;  
Best Local Similarity 80.0%; Pred. No. 8.7e-16;  
Matches 36; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Oy 19 MSDDKPFICTAGCCQRTNEDHLAVHKHKHMTLKFGPARNSKI 63  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 1 MGDDRXFCVNAFGCCQRTNEDHLAVHKHKHMTLKFGPARTDSV 45

RESULT 3  
US-09-503-954A-20  
Sequence 20, Application US/09503954A  
Patent No. 6610820  
GENERAL INFORMATION:  
APPLICANT: Bonny, Christophe  
TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL  
FILE REFERENCE: 20349-501  
CURRENT APPLICATION NUMBER: US/09/503,954A  
CURRENT FILING DATE: 2000-02-14  
PRIOR APPLICATION NUMBER: USSN 60/158,774  
PRIOR FILING DATE: 1999-10-12  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 20  
LENGTH: 29  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically synthesized  
US-09-503-954A-20

Query Match 17.7%; Score 137; DB 2; Length 29;  
Best Local Similarity 88.9%; Pred. No. 5e-08;  
Matches 24; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 37 TNEDHLAVHKHKHMTLKFGPARNSKI 63  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 1 TNEDHLAVHKHKHMTLKFGPARNSV 27

RESULT 4  
US-09-970-515A-20  
Sequence 20, Application US/09970515A  
Patent No. 6780970  
GENERAL INFORMATION:  
APPLICANT: Bonny, Christophe  
TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL  
FILE REFERENCE: 20349-501 DIV

GENERAL INFORMATION:  
CURRENT APPLICATION NUMBER: US/09/970,515A  
CURRENT FILING DATE: 2001-10-03  
PRIOR APPLICATION NUMBER: 09/503,954  
PRIOR FILING DATE: 2000-02-14  
PRIOR APPLICATION NUMBER: 60/158,774  
PRIOR FILING DATE: 1999-10-12  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 20  
LENGTH: 29  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically synthesized  
US-09-970-515A-20

Query Match 17.7%; Score 137; DB 2; Length 29;  
Best Local Similarity 88.9%; Pred. No. 5e-08;  
Matches 24; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 37 TNEDHLAVHKHKHMTLKFGPARNSKI 63  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 1 TNEDHLAVHKHKHMTLKFGPARNSV 27

RESULT 5  
US-08-819-177-6  
Sequence 6, Application US/08819177  
Patent No. 6043083  
GENERAL INFORMATION:  
APPLICANT: Davis, Roger J.  
APPLICANT: Dickens, Martin  
TITLE OF INVENTION: INHIBITORS OF THE JNK SIGNAL  
TITLE OF INVENTION: TRANSDUCTION PATHWAY AND METHODS OF USE  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/819,177  
FILING DATE: 28 April 1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. Peter  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 07917/037001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
OTHER INFORMATION: ATF-2 JNK-binding domain  
US-08-819-177-6

Query Match 14.8%; Score 115; DB 2; Length 27;  
Best Local Similarity 87.0%; Pred. No. 1.1e-05;  
Matches 20; Conservative 1; Mismatches 2; Indels 0; Gaps 0;



OTHER INFORMATION: Description of Artificial Sequence: designed ZFP  
US-09-941-450-15

Query Match 13.7%; Score 106.5; DB 2; Length 97;  
Best Local Similarity 39.7%; Pred. No. 0.00053;  
Matches 25; Conservative 9; Mismatches 24; Indels 5; Gaps 2;  
QY 17 WMSDDKPLCTAPCGGQRTNEDHLAVHKHCHMTLKG----PARNSKIEEPSVWETT 72  
DB 31 WH-TGERPFMCTSYCGKRFNTRDLARHRTHTGEEKFACPCPKRFRMSDHLKHIKT 89  
QY 73 HQD 75  
DB 90 HQN 92

RESULT 10  
US-09-949-016-11145  
Sequence 11145, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CLO01307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11145  
LENGTH: 522  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-11145

Query Match 13.3%; Score 103; DB 2; Length 522;  
Best Local Similarity 25.7%; Pred. No. 0.013;  
Matches 44; Conservative 18; Mismatches 63; Indels 46; Gaps 8;  
QY 9 SARQYKDLWM-SDDKPLCTAPCGGQRTNED----HLAVKH-----KHMT 52  
DB 305 SATNKNHVRHITGKPYVCTVPGCGKRFTEYSSLYKHHVHTCKPYTCSTCGKTYRQT 364  
QY 53 LKFGPARNSKIEEPSVWETHO-----DSPLD-HPE---STTSDEKEVP 92  
DB 365 STLAMHKSANHGELEATESEALYEQOLEAASAAESPPPRPRIAYLSEVKERDDI 424  
QY 93 LAQTAQPSAIVRPASLOVPNVLLTSSDSSVLIQAVPSTSTVITQAPS 143  
DB 425 PAQVAMVTE-----EDGAPQALITQDGA---QQVTIITSGAVVAEDSS 465

RESULT 11  
US-08-224-482-2  
Sequence 2, Application US/08224482  
Patent No. 5837692  
GENERAL INFORMATION:  
APPLICANT: Mercola, Dan  
TITLE OF INVENTION: Inhibition of the Mitogenic Activity of  
ADAMSON, Eileen D.  
TITLE OF INVENTION: PDGF by Mammalian EGR  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California

COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/224,482  
FILING DATE: 07-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-ME 9913  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 496 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-224-482-2

Query Match 13.2%; Score 102.5; DB 1; Length 496;  
Best Local Similarity 26.7%; Pred. No. 0.014;  
Matches 35; Conservative 22; Mismatches 51; Indels 23; Gaps 5;  
QY 22 DKPFLCTAPCGGQRTNED-----HLAVKHCHMTLKGPARNSKIEEPSVWETH 73  
DB 354 EKFPACDI--CGRKPARSDERKRHTKIHLRQKDKKADKSVVASPAASLSYSPVATSY 411  
QY 74 QDSPLPHESTSDKEVPLAQTAPQPSAIVRPASLOVPNVLLTSSDSSVLIQAVPST 133  
DB 412 -----PSP-ATTSPSPVPTSYSGSGSTYSPAHSGPSPSVATTFASV-----PPA 458  
QY 134 SSTVITQAPSS 144  
DB 459 FPTQVSSPFA 469

RESULT 12  
US-08-040-548-1  
Sequence 1, Application US/08040548  
Patent No. 5763209  
GENERAL INFORMATION:  
APPLICANT: Sukhatme, Vikas P.  
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE  
FUNCTIONAL DOMAINS OF DNA BINDING PROTEINS  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: 321 No. 5763209th Clark Street, Suite 800  
CITY: Chicago  
STATE: Illinois  
COUNTRY: U.S.A.  
ZIP: 60610  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/040,548  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Coughlin, Daniel F.  
REGISTRATION NUMBER: 36,111  
REFERENCE/DOCKET NUMBER: arcd067  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 744-0090  
TELEFAX: (312) 245-4961  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 533 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-040-548-1

Query Match 13.2%; Score 102.5; DB 1; Length 533;  
Best Local Similarity 26.7%; Pred. No. 0.016;  
Matches 35; Conservative 22; Mismatches 51; Indels 23; Gaps 5;

QY 22 DKPFLCTAPGCGQRTNED-----HLAVHKHKHMTLKFGRNRSKIEEPSVWETH 73  
Db 391 EKPFACDI--CGRKFARSDERKRHTKIHLRQKDKADKSVVASPAASLSLSPSVATSY 448  
QY 74 QDSPLPHPESTTSDEKEVPLAQTAQTSIAIVRPASLQVNPVLLTSSDSSVIIQAVPSPT 133  
Db 449 -----PSP-ATTSPPSPVPTSYSPGSGSTYSPAHSGFPSPSVATTFAV-----PPA 495  
QY 134 SSTVITQAPSS 144  
Db 496 PPTQVSSFPFA 506

RESULT 13  
US-08-466-344-1  
; Sequence 1, Application US/08466344  
; Patent No. 5773583  
; GENERAL INFORMATION:  
; APPLICANT: Sukhatme, Vikas P.  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO THE  
; TITLE OF INVENTION: FUNCTIONAL DOMAINS OF DNA BINDING PROTEINS  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: 321 No. 5773583th Clark Street, Suite 800  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 60610  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/466,344  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/040,548  
; FILING DATE: 31-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coughlin, Daniel F.  
; REGISTRATION NUMBER: 36,111  
; REFERENCE/DOCKET NUMBER: arcd067  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 744-0090  
; TELEFAX: (312) 245-4961  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 533 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-466-344-1

Query Match 13.2%; Score 102.5; DB 1; Length 533;

Best Local Similarity 26.7%; Pred. No. 0.016;  
Matches 35; Conservative 22; Mismatches 51; Indels 23; Gaps 5;  
QY 22 DKPFLCTAPGCGQRTNED-----HLAVHKHKHMTLKFGRNRSKIEEPSVWETH 73  
Db 391 EKPFACDI--CGRKFARSDERKRHTKIHLRQKDKADKSVVASPAASLSLSPSVATSY 448  
QY 74 QDSPLPHPESTTSDEKEVPLAQTAQTSIAIVRPASLQVNPVLLTSSDSSVIIQAVPSPT 133  
Db 449 -----PSP-ATTSPPSPVPTSYSPGSGSTYSPAHSGFPSPSVATTFAV-----PPA 495  
QY 134 SSTVITQAPSS 144  
Db 496 PPTQVSSFPFA 506

RESULT 14  
US-09-546-013-21  
; Sequence 21, Application US/09546013  
; Patent No. 6610504  
; GENERAL INFORMATION:  
; APPLICANT: Yuan, Chong-Shen  
; TITLE OF INVENTION: METHODS FOR ASSAYING S-ADENOSYLMETHIONINE-DEPENDENT Methyltransf  
; FILE REFERENCE: 10937-1652  
; CURRENT APPLICATION NUMBER: US/09/546,013  
; CURRENT FILING DATE: 2000-04-10  
; EARLIER APPLICATION NUMBER: 09/347,878  
; EARLIER FILING DATE: 1999-07-06  
; EARLIER APPLICATION NUMBER: 09/457,205  
; EARLIER FILING DATE: 1999-12-06  
; NUMBER OF SEQ ID NOS: 94  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 21  
; LENGTH: 533  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Human Egr Dna binding protein: zinc finger domain  
; PUBLIC INFORMATION:  
; PATENT DOCUMENT NUMBER: 5,866,325  
; PATENT FILING DATE: 1999-06-06  
; PUBLICATION DATE: 1999-02-02  
US-09-546-013-21

Query Match 13.2%; Score 102.5; DB 2; Length 533;  
Best Local Similarity 26.7%; Pred. No. 0.016;  
Matches 35; Conservative 22; Mismatches 51; Indels 23; Gaps 5;  
QY 22 DKPFLCTAPGCGQRTNED-----HLAVHKHKHMTLKFGRNRSKIEEPSVWETH 73  
Db 391 EKPFACDI--CGRKFARSDERKRHTKIHLRQKDKADKSVVASPAASLSLSPSVATSY 448  
QY 74 QDSPLPHPESTTSDEKEVPLAQTAQTSIAIVRPASLQVNPVLLTSSDSSVIIQAVPSPT 133  
Db 449 -----PSP-ATTSPPSPVPTSYSPGSGSTYSPAHSGFPSPSVATTFAV-----PPA 495  
QY 134 SSTVITQAPSS 144  
Db 496 PPTQVSSFPFA 506

RESULT 15  
5206152-2  
; Patent No. 5206152  
; APPLICANT: SUKHATME, VIKAS P.  
; TITLE OF INVENTION: CLONING AND EXPRESSION OF EARLY GROWTH  
; REGULATORY PROTEIN GENES  
; NUMBER OF SEQUENCES: 18  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/249,584  
; FILING DATE: 26-SEP-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 179,587

```

; FILLING DATE: 08-APR-1988
; SEQ ID NO:2:
; LENGTH: 533
5206152-2

Query Match      13.2%; Score 102.5; DB 6; Length 533;
Best Local Similarity 26.7%; Pred. No. 0.016;
Matches 35; Conservative 22; Mismatches 51; Indels 23; Gaps 5;

QY 22 DKPFLCTAPGCGQRTNED-----HLAVHKHGHMTLKFGPARNSKIEEPSVVETTH 73
Db 391 EKPFCADI--CGKRFARSDERKRHTKIHLRQDKADKSWASPAASSLSYSPVATSY 448
QY 74 QDSPLHPESWTTSDEKEVPLAQTAQPTSAIVRPASLOVPNVLLTSSDSSVIIQQAVPSP 133
Db 449 -----PSP-ATTFPSPVPTSYSSGSGSTYPSGHSFPSPSVATTFASV-----PPA 495

QY 134 SSTVITQAPSS 144
Db 496 FTTQVSSFPESA 506
```

Search completed: January 10, 2006, 23:07:30  
Job time : 24 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame plus\_p2n\_model.v8  
Run on: January 12, 2006, 17:52:40 ; Search time 3044 Seconds  
(without alignments)  
298.783 Million cell updates/sec

Title: AAH26175\_COPY\_50\_65  
Perfect score: 81  
Sequence: 1 emtkfgparnskiee 16

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5893141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/RAWLINGS10076905/runat\_10012006\_144823\_5320/app\_query.fasta\_1.199  
-DB=GenEmbl -QMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=RAWLINGS10076905@CGN\_1\_1\_4939@runat\_10012006\_144823\_5320 -NCPU=6  
-ICPU=3 -NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:  
1: gb.ba.\*  
2: gb.in.\*  
3: gb.env.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pr.\*  
9: gb.ro.\*  
10: gb.sts.\*  
11: gb.sv.\*  
12: gb.un.\*  
13: gb.vi.\*  
14: gb.htg.\*  
15: gb.pl.\*

RESULT 1  
RATRATF2  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	75.3	1209	9	RATRATF2 m
2	61	75.3	1302	9	S76659
3	61	75.3	1370	8	BC026175 Homo sapi

4	61	75.3	1464	9	AF483482	AF483482 Mus muscu
5	61	75.3	1464	9	AF483483	AF483483 Mus muscu
6	61	75.3	1520	5	XLJ16158	U16158 Xenopus lae
7	61	75.3	1596	9	S76657	S76657 cyclic AMP
8	61	75.3	1647	6	AX821940	AX821940 Sequence
9	61	75.3	1647	6	AX822000	AX822000 Sequence
10	61	75.3	1647	8	HSCREBP1	X15875 Human mRNA
11	61	75.3	1768	8	DQ003047	DQ003047 Homo sapi
12	61	75.3	1812	5	GGN17724	Y17724 Gallus gall
13	61	75.3	1843	8	DQ003049	DQ003049 Homo sapi
14	61	75.3	1852	9	RNU38938	U38938 Rattus norv
15	61	75.3	1883	8	DQ003037	DQ003037 Homo sapi
16	61	75.3	1901	8	AB169461	AB169461 Macaca fa
17	61	75.3	1918	8	DQ003041	DQ003041 Homo sapi
18	61	75.3	1957	8	DQ003039	DQ003039 Homo sapi
19	61	75.3	1965	8	DQ003044	DQ003044 Homo sapi
20	61	75.3	2080	8	DQ003043	DQ003043 Homo sapi
21	61	75.3	2112	8	DQ003036	DQ003036 Homo sapi
22	61	75.3	3909	9	BC082596	BC082596 Mus muscu
23	61	75.3	3957	9	BC042210	BC042210 Mus muscu
24	61	75.3	4851	11	AF050498	AF050498 Fusion tr
25	61	75.3	11797	8	AC074291	AC074291 Homo sapi
26	61	75.3	166766	14	AC119682	AC119682 Rattus no
27	61	75.3	173793	14	AC129355	AC129355 Rattus no
28	61	75.3	176033	14	AL161794	AL161794 Homo sapi
29	61	75.3	178342	8	AC007435	AC007435 Homo sapi
30	61	75.3	207639	9	AL844581	AL844581 Mouse DNA
31	57	70.4	501	6	AX321198	AX321198 Sequence
32	55	67.9	229	6	BD026693	BD026693 Sequence
33	55	67.9	229	6	AX887083	AX887083 Sequence
34	55	67.9	787	8	BC042363	BC042363 Homo sapi
35	55	67.9	800	6	CO491274	CO491274 Sequence
36	55	67.9	800	6	CO497144	CO497144 Sequence
37	55	67.9	1227	6	CQ720976	CQ720976 Sequence
38	55	67.9	1485	8	HSATFAL	X57197 H.sapiens m
39	55	67.9	2423	5	AY28940	AY28940 Danio rer
40	55	67.9	2755	6	CO490766	CO490766 Sequence
41	55	67.9	2755	6	CO496609	CO496609 Sequence
42	55	67.9	2758	8	HSATFA	X52943 Human mRNA
43	55	67.9	2867	9	BC026483	BC026483 Mus muscu
44	55	67.9	3334	5	CR942576	CR942576 Xenopus t
45	55	67.9	113038	14	CR933816	CR933816 Danio rer

## ALIGNMENTS

RATRATF2	Rat RATF2 mRNA, complete cds.	1209 bp	mRNA	linear	ROD 27-APR-1993
M65148					
M65148.1	GI:206569				
Rattus norvegicus (Norway rat)					
Rattus norvegicus					
Eukaryota; Metazoa;					
Mammalia; Eutheria;					
Sciurognathi; Muroidae;					
1 (bases 1 to 1209)					
Kageyama,R., Sasai,Y. and Nakanishi,S.					
Molecular characterization of transcription factors that bind to					
the cAMP responsive region of the substance P precursor gene. cDNA					
cloning of a novel C/EBP-related factor					
J. Biol. Chem. 266 (23), 15525-15531 (1991)					
source text: Rat, cDNA to mRNA.					
Location/Qualifiers					
1..1209					
/organism="Rattus norvegicus"					
/mol_type="mRNA"					
/strain="Sprague-Dawley"					
/db_xref="taxon:10116"					
/dev_stage="adult"					

```

gene
1. .1209
/ gene="RATP2"
CDS
28. .1197
/ gene="RATP2"
/codon_start=1
/protein_id="AA442013.1"
/db_xref="GI:206570"
/translation="MSDDKPFICTAPGCGQRTNEDHLAVHKHMTLKFGPARNDS
VIVAOQTPTTFLKNCSEVGLFNLASPFENFKASDEDDIKWPLDLSPLATPIIR
SKIEPSVVEITHQDSPLPHPSSTNDEKLVRPVTMVPSPVGIPIGPSGPQVQSEAKM
RLKAALTOQHPVTNGDTVKGHSGLVRAQSEESRQSLQQPATSTTETPASPAHTTP
OTQNTSGRRRAANEDPKRRKFLERNRAAARCRQKRWVQSLKKAEDLSLNG
OLQSEVTLRNVAOLKLLAHKDCPVTAMOKKSGYHTADKDDSDSLSPSSPHT
AIQHSVSSTNSGVSSTKTEAGNTSVLTQADQSTEPALSQIVMAPSSQAQPSGS"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 0.029 Length: 1209
Score: 61.00 Matches: 11
Percent Similarity: 85.71% Conservative: 1
Best Local Similarity: 78.57% Mismatches: 2
Query Match: 75.31% Indels: 0
DB: 9 Gaps: 0

```

AAH26175\_COPY\_50\_65 (1-16) x RATP2 (1-1209)

```

QY 1 GluMetThrLeuLysPheGlyProAlaArgAenSerLysIle 14
|||||
DB 121 GAGATGACACTGAAATTTGGTCACGACGTAATGACAGTGTC 162
|||||

```

## RESULT 2

```

S76659 1302 bp mRNA linear ROD 07-MAY-1993
LOCUS cyclic AMP response element DNA-binding protein isoform 2
DEFINITION {alternative splicing product} [mice, EL4, mRNA, 1302 nt].

```

S76659.1 GI:243430

```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus sp.

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

```

1 (bases 1 to 1302)

```

AUTHORS Georgopoulos, K., Morgan, B.A. and Moore, D.D.
TITLE Functionally distinct isoforms of the CRE-BP DNA-binding protein
mediate activity of a T-cell-specific enhancer
Mol. Cell. Biol. 12 (2), 747-757 (1992)

```

1531087

```

REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gisbseq 76659] from the original journal article.
Location/Qualifiers

```

```

FEATURES
source
1. .1302
/organism="Mus sp."
/mol_type="mRNA"
/db_xref="taxon:10095"

```

## gene

## CDS

```

1. .1302
/ gene="cyclic AMP response element DNA-binding protein
isoform 2, CRE-BP2"
188. .1264
/ gene="cyclic AMP response element DNA-binding protein
isoform 2, CRE-BP2"
/codon_start=1
/product="cyclic AMP response element DNA-binding protein
isoform 2"
/notes="CRE-BP2"

```

/protein\_id="AAB21129.1"

```

/db_xref="GI:243431"
/translation="MTLTKGPARNDSVIVADQPTTTRFLKNCSEVGLFNLASPFEN
EFKASDEDDIKWPLDLSPLATPIIRSKIEPSVVEITHQDSPLPHPSSTNDEKLVR
PVTMVPSPVGIPIGPSGPQVQSEAKRLKAALTOQHPVTNGDTVKGHSGLVRAQSE
ESRQSLQQPATSTTETPASPAHTTPTQNTSGRRRAANEDPKRRKFLERNRAA
SRCRQKRWVQSLKKAEDLSLNGQLQSEVTLRNVAOLKLLAHKDCPVTAMQ"

```

```

KKSQYHTADKDDSDSLSPSSPHTAQHSVSSTNSGVSSTKAEAVATSVLTQAD
QSTEPALSQIVMAPSSQAQPSVD"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 0.0314 Length: 1302
Score: 61.00 Matches: 11
Percent Similarity: 85.71% Conservative: 1
Best Local Similarity: 78.57% Mismatches: 2
Query Match: 75.31% Indels: 0
DB: 9 Gaps: 0

```

AAH26175\_COPY\_50\_65 (1-16) x S76659 (1-1302)

```

QY 1 GluMetThrLeuLysPheGlyProAlaArgAenSerLysIle 14
|||||
DB 185 GAGATGACACTGAAATTTGGTCACGACGTAATGACAGTGTC 226
|||||

```

## RESULT 3

```

BC026175 1370 bp mRNA linear PRI 16-SEP-2003
LOCUS Homo sapiens activating transcription factor 2, mRNA (cdna clone
DEFINITION IMAGE:4308091), complete cds.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1370)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E.,
Schnerich, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

```

2 (bases 1 to 1370)

Straussberg, R.

Direct Submission

```

Submitted (02-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)

Tissue Procurement: ATCC/DC/DTP

CDNA Library Preparation: CLONTECH Laboratories, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLML)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>contact: [amadnan@systemsbio.org](mailto:amadnan@systemsbio.org)

Anup Madan, Jeessica Fahey, Erin Helton, Mark Kerteman, Anuradha

Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 38 Row: c Column: 20  
 This clone has the following problem: The cds is short compared to the longest cds in the locus.

## FEATURES

source

Location/Qualifiers  
 1..1370  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone\_image="IMAGE:4308091"  
 /tissue\_type="Skin, melanotic melanoma, high MDR."  
 /clone\_lib="NIH MGC 62"  
 /lab\_host="DH10B"  
 /note="Vector: pDNR-LIB"

gene

1..1370  
 /gene="ATF2"  
 /note="synonyms: CRE-BP1, TREB7, HB16"  
 /db\_xref="GeneID:1386"  
 /db\_xref="MIM:123811"

CDS

267..896  
 /gene="ATF2"  
 /codon\_start=1  
 /product="ATF2 protein"  
 /protein\_id="AAH26175.1"  
 /db\_xref="GI:20072897"  
 /db\_xref="GeneID:1386"  
 /db\_xref="MIM:123811"  
 /translation="MKFKLVHNSAQYKDLNMSDDKPLCTAPCGQRTNEDHLAV  
 HKHHEMTLKFGPARNSIVADQPTPTFLKNCSEVGLFNELASPFENFKASEDDIKGMPLDLSPLATPIR  
 DIKNPLDLSPLATPIIRSKIEPSVETTHQDSPLPHPESTTSDEKEVPLAQTAQPT  
 SAIVRPASLQVNVLTSSDSSVIIQAQVPSSTSTVITQAPSSNRPIV"

## ORIGIN

Alignment Scores:  
 Pred. No.: 0.0331 Length: 1370  
 Score: 61.00 Matches: 11  
 Percent Similarity: 85.71% Conservative: 1  
 Best Local Similarity: 78.57% Mismatches: 2  
 Query Match: 75.31% Indels: 0  
 DB: 8 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x BC026175 (1-1370)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAenSerLysIle 14  
 |||||  
 DB 414 GAGATGACACTGAAATTTGGTCCAGCAGTAATGACAGTGC 455

RESULT 4

AF483482

LOCUS AF483482 1464 bp mRNA linear ROD 19-MAR-2002  
 DEFINITION Mus musculus strain ILS CAMP response element binding protein 1

ACCESSION

AF483482

VERSION

AF483482.1 GI:19548717

KEYWORDS

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1464)

Ehringer, M.A., Thompson, J., Conroy, O., Xu, Y., Yang, F., Canniff, J.,

Beeson, M., Gordon, L., Bennett, B., Johnson, T.E. and Sikela, J.M.

High-throughput sequence identification of gene coding variants

within alcohol-related QTLs

Mamm. Genome 12 (8), 657-663 (2001)

11471062

2 (bases 1 to 1464)

Ehringer, M.A., Thompson, J., Conroy, O., Xu, Y., Yang, F., Canniff, J.,

Beeson, M., Gordon, L., Bennett, B., Johnson, T.E. and Sikela, J.M.

Direct Submission

Submitted (14-FEB-2002) Pharmacology, University of Colorado Health

## FEATURES

source

Location/Qualifiers  
 1..1464  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="ILS"  
 /db\_xref="taxon:10090"

CDS

1..1464  
 /codon\_start=1  
 /product="CAMP response element binding protein 1"  
 /protein\_id="AAL90756.1"  
 /db\_xref="GI:19548718"  
 /translation="MSDDKPLCTAPCGQRTNEDHLAVHKHHEMTLKFGPARNS  
 VIVADQPTPTFLKNCSEVGLFNELASPFENFKASEDDIKGMPLDLSPLATPIR  
 SKIEPSVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVNVLTSS  
 SDSSVIIQAQVPSSTSTVITQAPSSNRPIVVPVPGFPPLLLHLPNGQMPVAPASIT  
 SSNVHVPALVELRPVTVNVPVPGIPGSPSSQPVQSEAKMRLLKALTOQHPVTVNGDT  
 VKGHGPAALVTSSESRPQSQQPAITSTETPASPAPHTPTQNTSGRRRAANEDPD  
 EKRRFLERNRAAARCRQKRWVQSLKABDLSLNGQLQSEVTLRLNVAQLKQ  
 LLLAKDCPVTAMQKSGYHTADKDDSSDLSVPSPHTEATIQHSVSTSGVSSSTSK  
 AEAVATSVLTQADQSTEPALSQIVMAPPSQAPSGS"

## ORIGIN

Alignment Scores:  
 Pred. No.: 0.0356 Length: 1464  
 Score: 61.00 Matches: 11  
 Percent Similarity: 85.71% Conservative: 1  
 Best Local Similarity: 78.57% Mismatches: 2  
 Query Match: 75.31% Indels: 0  
 DB: 9 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x AF483482 (1-1464)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAenSerLysIle 14

|||||  
 DB 94 GAGATGACACTGAAATTTGGTCCAGCAGTAATGACAGTGC 135

RESULT 5

AF483483

LOCUS AF483483 1464 bp mRNA linear ROD 19-MAR-2002  
 DEFINITION Mus musculus strain ISS CAMP response element binding protein 1

ACCESSION

AF483483

VERSION

AF483483.1 GI:19548719

KEYWORDS

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1464)

Ehringer, M.A., Thompson, J., Conroy, O., Xu, Y., Yang, F., Canniff, J.,

Beeson, M., Gordon, L., Bennett, B., Johnson, T.E. and Sikela, J.M.

High-throughput sequence identification of gene coding variants

within alcohol-related QTLs

Mamm. Genome 12 (8), 657-663 (2001)

11471062

2 (bases 1 to 1464)

Ehringer, M.A., Thompson, J., Conroy, O., Xu, Y., Yang, F., Canniff, J.,

Beeson, M., Gordon, L., Bennett, B., Johnson, T.E. and Sikela, J.M.

Direct Submission

Submitted (14-FEB-2002) Pharmacology, University of Colorado Health

Science Center, 4200 E. 9th Ave C236, Denver, CO 80262, USA

Location/Qualifiers

1..1464

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="ISS"

/db\_xref="taxon:10090"

1..1464

/codon\_start=1

/product="CAMP response element binding protein 1"

/protein\_id="AAL90757.1"

```
/db_xref="GI:19548720"
/translation="MSDDKPEFLCTAPCGORFTNEDHLAVHKHKEHTLTKFGPARNDS
VIVADQTPITRFLKCEEVGLNELASPFENEFKASDEDDIKKPLDLSPLATPIIR
SKIEEPSVETTHQDSPLPHEPSTTSDEKEVPLAQTAQTSALVRPASLQVNPVLITS
SDSVIIQQAQVSPSTVITQAPSNRPVVPVPGFPLLLHLPLNGQTMVPAIPASIT
SSNVHPAAVPLVRPVTMVPSPVPGIPGSPQPVQSEAKRWLKAALTOQHPPVTNGDT
VKHGSGLVRTQSESRPOSLOQPATSTTETPASPAHTTPTQNTSGRRRAANEDPD
EKERKFLERNRAASRCRQKRVVQSLKKAEDLSLNGLOLQSEVTLRNEVAOLKQ
LLLAHKDCPVTAMQKSGYHADKDDSDSDLSVPSPPTAEIOHSSVSTNSGVSTSK
AEAVATSVLTQMAQTEPALSQIVMAPSSQHPSGR"
```

## ORIGIN

Alignment Scores:

Pred. No.:	0.0356	Length:	1464
Score:	61.00	Matches:	11
Percent Similarity:	85.71%	Conservative:	1
Best Local Similarity:	78.57%	Mismatches:	2
Query Match:	75.31%	Indels:	0
DB:	9	Gaps:	0

AAH26175\_COPY\_50\_65 (1-16) x AF483483 (1-1464)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14  
|||||  
Db 94 GAGATGACACTGAATTTGGTCCAGCAGCGTAATGACAGTGTC 135  
|||||

## RESULT 6

XLU16158 1520 bp mRNA linear VRT 04-JUL-1995  
LOCUS Xenopus laevis activating transcription factor 2 (ATF2) mRNA,  
DEFINITION complete cds.

ACCESSION U16158

VERSION U16158.1

KEYWORDS GI:887779

SOURCE Xenopus laevis (African clawed frog)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
Xenopodinae; Xenopus.

## REFERENCE

1 Villareal,X.C. and Richter,J.D.  
AUTHORS Analysis of ATF2 gene expression during early Xenopus laevis  
development  
JOURNAL Gene 153 (2), 225-229 (1995)  
PUBMED 7875593  
REFERENCE 2 (bases 1 to 1520)  
AUTHORS Richter,J.D.  
TITLE Direct Submission  
JOURNAL Submitted (20-OCT-1994) Joel D. Richter, Worcester Foundation for  
Experimental Biology, 222 Maple Avenue, Shrewsbury, MA 01545, USA  
COMMENT On Jul 6, 1995 this sequence version replaced gi:710325.

## FEATURES

source  
1..1520  
Location/Qualifiers  
1..1520  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="pATF2"  
/cell\_type="oocyte"  
1..1520  
/gene="ATF2"  
25..1485  
/gene="ATF2"  
/codon\_start=1  
/evidence=experimental  
/product="activating transcription factor 2"  
/protein\_id="AA69518.1"  
/db\_xref="GI:710326"  
translation="MSDDKPEFLCTAPCGORFTNKHDLAVHKHKEHTLTKFGPARNDS  
VIVADQTPITRFLKCEEVGLNELASPFENEFKASDEDDIKKPLDLSPLATPIIR  
SKIEEPSVETTHQDSPLPHEPSTTSQHNQVPLAQTAQTSALVRPASLQVNPVLITS  
SDSVIIQQAQVSPSTVITQAPSNRPVVPVPGFPLLLHLPLNGQTMVPAIPASIT  
SSNVHPAAVPLVRPVTMVPSPVPGIPGSPQPVQSEAKRWLKAALTOQHPPVTNGDT  
VKHGSGFALTQSESRPOSLOQPATSTTETPASPAHTTPTQNTSGRRRAANEDPD

## gene

## CDS

```
EkRSKIQRNRAAASRCRQKRVVQSLKKADELISLNGYLQNEVTLRNEVAOLKQ  
LLLAHKDCPVTAMQKSGYHADKDDSDSDISVSPSPPTAEIOHSSVSTNSGVSTSK  
AEAVATSVLTMAHRTPEALSQIVMAPSSQHPSGR"
```

## polyA\_site

/gene="ATF2"

/note="18 A nucleotides"

## ORIGIN

Alignment Scores:

Pred. No.:	0.037	Length:	1520
Score:	61.00	Matches:	11
Percent Similarity:	85.71%	Conservative:	1
Best Local Similarity:	78.57%	Mismatches:	2
Query Match:	75.31%	Indels:	0
DB:	5	Gaps:	0

AAH26175\_COPY\_50\_65 (1-16) x XLU16158 (1-1520)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14  
|||||  
Db 118 GAGATGACACTGAATTTGGTCCAGCAGCGTAATGACAGTGTC 159  
|||||

## RESULT 7

S76657

LOCUS

DEFINITION

cyclic AMP response element DNA-binding protein isoform 1

{alternative splicing product} (mice, EL4, mRNA, 1596 nt).

ACCESSION S76657

VERSION S76657.1

KEYWORDS GI:243428

SOURCE Mus sp.

ORGANISM

Mus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1596)

AUTHORS Georgopoulos,K., Morgan,B.A. and Moore,D.D.

TITLE Functionally distinct isoforms of the CRE-BP DNA-binding protein

mediate activity of a T-cell-specific enhancer

JOURNAL Mol. Cell. Biol. 12 (2), 747-757 (1992)

PUBMED 1531087

REMARK GenBank staff at the National Library of Medicine created this

entry [NCBI gibseq 76657] from the original journal article.

FEATURES

Location/Qualifiers

1..1596

/organism="Mus sp."

/mol\_type="mRNA"

/db\_xref="taxon:10095"

1..1596

/gene="cyclic AMP response element DNA-binding protein

isoform 1, CRE-BP1"

188..1558

/gene="cyclic AMP response element DNA-binding protein

isoform 1, CRE-BP1"

/note="CRE-BP1"

/codon\_start=1

/product="cyclic AMP response element DNA-binding protein

isoform 1"

/protein\_id="AA821128.1"

/db\_xref="GI:243429"

/translation="MTLKFGPARNDSVIVADQTPITRFLKCEEVGLNELASPFEN

EFKASDEDDIKKPLDLSPLATPIIRSKIEEPSVETTHQDSPLPHEPSTTSDEKEVP

LAQTAQTSALVRPASLQVNPVLITSDDSVIIQQAQVSPSTVITQAPSNRPVVP

VPGFPLLLHLPLNGQTMVPAIPASITSSNVHPAAVPLVRPVTMVPSPVPGIPGSPQ

PVQSEAKRWLKAALTOQHPPVTNGDTVKHGSGLVRTQSESRPOSLOQPATSTTETP

ASPAHTTPTQNTSGRRRAANEDPDKRKRLERNRAASRCRQKRVVQSLKKA

EDLSLNGLOLQSEVTLRNEVAOLKQLLAHDKDDSDISVSPSPPTAEIOHSSVSTNS

GVSTSKAEAVATSVLTQMAQTEPALSQIVMAPSSQHPSGR"

## ORIGIN

Alignment Scores:

Pred. No.: 0.039

Length: 1596

Score: 61.00 Matches: 11  
 Percent Similarity: 85.71% Conservative: 1  
 Best Local Similarity: 78.57% Mismatches: 2  
 Query Match: 75.31% Indels: 0  
 DB: 9 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x S76657 (1-1596)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14  
 DB 185 GAGATGACACTGAATTTGGTCCAGCACGTAATGACAGTGTC 226

RESULT 8  
 LOCUS AX821940 1647 bp DNA linear PAT 10-DEC-2003  
 DEFINITION Sequence 68 from Patent WO03068961.  
 ACCESSION AX821940  
 VERSION AX821940.1 GI:39725161  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.

REFERENCE 1  
 AUTHORS Andrews, P.A., Walsh, J.A. and Gokhale, P.A.  
 TITLE Method to modify differentiation of pluripotential stem cells  
 JOURNAL Patent: WO 03068961-A 68 21-AUG-2003;  
 Axordia Limited (GB)  
 FEATURES Location/Qualifiers  
 source  
 1..1647  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

## ORIGIN

Alignment Scores:  
 Pred. No.: 0.0403 Length: 1647  
 Score: 61.00 Matches: 11  
 Percent Similarity: 85.71% Conservative: 1  
 Best Local Similarity: 78.57% Mismatches: 2  
 Query Match: 75.31% Indels: 0  
 DB: 9 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x AX821940 (1-1647)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14  
 DB 174 GAGATGACACTGAATTTGGTCCAGCACGTAATGACAGTGTC 215

RESULT 9  
 LOCUS AX822000 1647 bp DNA linear PAT 10-DEC-2003  
 DEFINITION Sequence 128 from Patent WO03068961.  
 ACCESSION AX822000  
 VERSION AX822000.1 GI:39725221  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.

REFERENCE 1  
 AUTHORS Andrews, P.A., Walsh, J.A. and Gokhale, P.A.  
 TITLE Method to modify differentiation of pluripotential stem cells  
 JOURNAL Patent: WO 03068961-A 128 21-AUG-2003;  
 Axordia Limited (GB)  
 FEATURES Location/Qualifiers  
 source  
 1..1647  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

## ORIGIN

## Alignment Scores:

Pred. No.: 0.0403 Length: 1647  
 Score: 61.00 Matches: 11  
 Percent Similarity: 85.71% Conservative: 1  
 Best Local Similarity: 78.57% Mismatches: 2  
 Query Match: 75.31% Indels: 0  
 DB: 6 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x AX822000 (1-1647)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14  
 DB 174 GAGATGACACTGAATTTGGTCCAGCACGTAATGACAGTGTC 215

RESULT 10  
 LOCUS HSCREBP1 1647 bp mRNA linear PRI 12-SEP-1993  
 DEFINITION Human mRNA for cAMP response element (CRE-BP1) binding protein.  
 ACCESSION X15875  
 VERSION X15875.1 GI:30214  
 KEYWORDS CREBP1 gene; DNA binding protein; leucine zipper; transcription factor.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.

REFERENCE 1 (bases 1 to 1647)

AUTHORS Maekawa, T., Sakurai, H., Kamei-Ishii, C., Sudo, T., Yoshimura, T.,  
 Fujisawa, J., Yoshida, M. and Ishii, S.  
 TITLE Leucine zipper structure of the protein CRE-BP1 binding to the  
 cyclic AMP response element in brain

JOURNAL EMBO J. 8 (7), 2023-2028 (1989)

PMID 2529117

COMMENT Data kindly reviewed (20-FEB-1990) by Ishii S.

FEATURES Location/Qualifiers

## source

1..1647  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="brain"  
 /clone\_lib="lambda gcl1"  
 /dev\_stage="foetus"  
 27..1544  
 /note="unnamed protein product; cAMP response element  
 binding protein (AA 1-505)"  
 /codon\_start=1  
 /protein\_id="CAA33886.1"  
 /db\_xref="GI:30215"  
 /db\_xref="GOA:P15336"  
 /db\_xref="UniProt/Swiss-Prot:P15336"

## CDS

/translation="MKFKLVNSARQYKDLNMSDDKPFLLTAPCGGRTNEDHLAV  
 HKHGHETLKPGRNDSIVADQTPTRTKNCEVGLFNEELASPEFNEFKGASED  
 DIKKMPLDPLATPIIRSKIEPSVVTTHQDSPLPHPESTTDEKEVPLAQTAQPT  
 SAIVRPSLQVNPVLLTSSDSSVITQAVPSPTSTVITQAPSSNRPTVPVPGPFLL  
 LHLPSGOTMPVAIPASITSSNVHPAAVPLVRVPTVPSVPGIPQSPSPQVQSEAKM  
 RLKALAQHPPTVNGDTVKGHSGLVRTQSEESRPOSLOQPATSTTETPASHTTP  
 QTQTSRRRRANEDDEKRFELERNRAAASRCRQKRVWVQSLKKAEDLSLNG  
 QLQSEVTLRNREVAQLKLLAHKDCPVTAQMKSGYHTADKDSSEDISVPSSPHTE  
 AIQHSVSVTNGSVSTSKAEAVATSVLTQADQSTEPALSIQVMAVPSQSPGSG"

## ORIGIN

Alignment Scores:  
 Pred. No.: 0.0403 Length: 1647  
 Score: 61.00 Matches: 11  
 Percent Similarity: 85.71% Conservative: 1  
 Best Local Similarity: 78.57% Mismatches: 2  
 Query Match: 75.31% Indels: 0  
 DB: 8 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x HSCREBP1 (1-1647)





CDS

```
317. .1780
/gene="ATF2"
/note="Contains 14 exons total; contains a nuclear
localization sequence; alternatively spliced"
/codon_start=1
/product="activating transcription factor 2 splice variant
ATF2-var1"
/protein_id="AAV17203.1"
/db_xref="GI:62866357"
/translation="MSDDKPFLLCTAPGCGQRTNEDHLAVHKHHEMTLKFGPARNDS
VIVADQTPTRFLKNCIEVGLNELASPFENEFKASEDDIKMPLDLSPLATPIIR
SKIEPSVETTHQSDPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVFNVLITS
SDSVIIQAVPSTSVITQAPSSNRPIVVPVPGFPLLLHLPSGOTMPVAIPASIT
SSNVHVPAAVPLVRVTVWVPSVPGIPGSSQPVQSEAKMRLKALTOCHPPVTNGDT
VKHGSGLVRTQSESRPQSLOQPATSTTETPASPAHTTPTQTSGRRRRAANEDPD
EKRRFLERNRAAAKRCQKRWVQSLKKAEDLSSLQSEVILLRNEVAQLKQ
LLLAHKDCPVTAMQKKSGYHTADKDDSDSDISVPSSPHTETAIQHSVSSTNGVSSTSK
AEAVATSVLTQADQSTBPALSQIVMAPSSQSPSGS"
```

ORIGIN

Alignment Scores:

Pred. No.:	0.0466	Length:	1883
Score:	61.00	Matches:	11
Percent Similarity:	85.71%	Conservative:	1
Best Local Similarity:	78.57%	Mismatches:	2
Query Match:	75.31%	Indels:	0
DB:	8	Gaps:	0

AAH26175\_COPY\_50\_65 (1-16) x DQ003037 (1-1883)

Qy	1	GluMetThrLeuLysPheGlyProAlaArgAanSerLysIle	14
Db	410	GAGATGACACTGAAATTTGGTCCAGCACGTAATGACAGTGTC	451

Search completed: January 12, 2006, 20:02:05  
Job time : 3049 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2006, 12:45:54 ; Search time 400 Seconds  
(without alignments)  
266.588 Million cell updates/sec

Title: Aah26175\_COPY\_50\_65  
Perfect score: 81  
Sequence: 1 emtlkfgparnekies 16

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-Q=/cgn2.1/USPTO.spool/RAWLINGS10076905/runat\_10012006\_144822\_5310/app\_query.fasta\_1.199  
-DB=N Geneseq -QMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-DOFWT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=RAWLINGS10076905 @CGN\_1\_727 @runat\_10012006\_144822\_5310 -NCPU=6  
-ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq\_21.\*

1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	75.3	1518	14 ADV42650	Adv42650 Human psy
2	61	75.3	1647	10 ADK67038	Adk67038 Gene #128
3	61	75.3	1647	10 ADK66978	Adk66978 Gene #68
4	61	75.3	1852	10 ADB53173	Adb53173 Primary r

5	58	71.6	51	10	ADC22243	Adc22243 Protein b
6	57	70.4	501	3	AAS61674	Aas61674 Lung emal
7	55	67.9	229	6	AAC02948	Aac02948 Human sec
8	55	67.9	800	5	ABV28993	Abv28993 Human pro
9	55	67.9	800	5	ABV23152	Abv23152 Human pro
10	55	67.9	1239	12	ADP80924	Adp80924 Mouse C13
11	55	67.9	2755	5	ABV28463	Abv28463 Human pro
12	55	67.9	2755	5	ABV22640	Abv22640 Human pro
13	55	67.9	2758	13	ACF87477	Acf87477 Human SIR
14	52	64.2	51	10	ADC22231	Adc22231 Protein b
15	46	56.8	397	5	ABV14146	Abv14146 Human pro
16	46	56.8	432	5	ABV44079	Abv44079 Human pro
17	46	56.8	432	5	ABV35240	Abv35240 Human pro
18	45	55.6	497	4	ABL03327	Ab103327 Drosophill
19	45	55.6	29604	2	AXA83005	Aax83005 Partial m
20	45	55.6	110000	2	AAV21209_14	Continuation (15 o
21	44	54.3	865	8	ABZ51307	Abz51307 Aspergill
22	44	54.3	110000	14	ADZ59507_2	Continuation (3 of
23	43	53.1	406	4	AA181886	Aa181886 Human pol
24	43	53.1	444	8	ACA29875	Aca29875 Prokaryot
25	43	53.1	694	6	ABQ19566	Abq19566 Oligonuel
26	43	53.1	694	6	ABQ19567	Abq19567 Oligonuel
27	43	53.1	803	6	ABQ19762	Abq19762 Oligonuel
28	43	53.1	803	6	ABQ19763	Abq19763 Oligonuel
29	43	53.1	1317	8	ACA21908	Aca21908 Prokaryot
30	43	53.1	1561	13	ADQ84003	Adq84003 Human tum
31	43	53.1	1561	13	ADQ86572	Adq86572 Human tum
32	43	53.1	1561	13	ADQ83403	Adq83403 Human tum
33	43	53.1	1854	12	ADO28895	Ado28895 CDNA enco
34	43	53.1	1980	1	AAN80919	Aan80919 Sequence
35	43	53.1	2000	11	ACL37790	Ac137790 Rice stre
36	43	53.1	2112	2	AAQ28370	Aaq28370 Herba-T c
37	43	53.1	2125	2	AAZ41316	Aaz41316 Human nor
38	43	53.1	2260	13	ADX52131	Adx52131 Plant ful
39	43	53.1	2309	12	ADH13166	Adh13166 Human mal
40	43	53.1	2309	14	AEA15053	Aea15053 Human pol
41	43	53.1	2373	13	ADX54078	Adx54078 Plant ful
42	43	53.1	2400	6	ABK34592	Abk34592 Human cdn
43	43	53.1	2422	14	ADX05487	Adx05487 Cyclin-de
44	43	53.1	2557	6	ABZ35333	Abz35333 Human gen
45	43	53.1	8974	4	ABL14100	Ab114100 Drosophill

#### ALIGNMENTS

RESULT 1  
ADV42650  
ID ADV42650 standard; cDNA; 1518 BP.  
XX  
AC ADV42650;  
XX  
DT 10-MAR-2005 (first entry)  
XX  
DE Human psychoneuroendocrinimmune expressed sequence tag SEQ ID NO 278.  
XX  
KW microarray; psychoneuroendocrinimmune; chronic fatigue;  
KW non-insulin dependent diabetes; allergy; immune disorder; inflammation;  
KW cancer; neoplasm; infection; expressed sequence tag; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2004108999-A2.  
XX  
PD 16-DEC-2004.  
XX  
PF 04-JUN-2004; 2004WO-US017686.  
XX  
PR 04-JUN-2003; 2003US-0475915P.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Nicholson A, Vernon SD;  
XX

DR WPI; 2005-031692/03.  
XX  
PT New microarray comprising probes for genes involved in  
PT psychoneuroendocrinimmune (PNI) activity, useful in diagnosing a  
PT condition associated with PNI activity, e.g., inflammatory or infectious  
PT diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 278; 254pp; English.  
XX  
CC The invention relates to a new microarray which comprises probes for  
CC genes involved in psychoneuroendocrinimmune (PNI) activity. The  
CC microarray is useful in diagnosing a condition associated with PNI  
CC activity, such as CFS, type-2 diabetes, allergic condition, inflammation,  
CC cancer and infection. The present sequence represents a  
CC psychoneuroendocrinimmune gene expressed sequence tag. Note the  
CC specificatio mentions SEQ ID NO of up to 3314 but only sequences up to  
CC SEQ ID NO 1829 are provided.  
XX  
XX Sequence 1518 BP; 461 A; 377 C; 316 G; 364 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 0.0628 Length: 1518  
Score: 61.00 Matches: 11  
Percent Similarity: 85.71% Conservative: 1  
Best Local Similarity: 78.57% Mismatches: 2  
Query Match: 75.31% Indels: 0  
DB: 14 Gaps: 0  
AAH26175\_COPY\_50\_65 (1-16) x ADV42650 (1-1518)  
QY 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14  
DB 148 GAGATGACACTGAAATTTGGTCCAGCAGCTAATGACAGTGTC 189  
RESULT 2  
ADK67038  
ID ADK67038 standard; DNA; 1647 BP.  
XX  
AC ADK67038;  
XX  
XX 06-MAY-2004 (first entry)  
DT  
DE Gene #128 for inhibitory RNA to manipulate stem cell phenotype.  
XX  
XX ds; gene; inhibitory RNA; RNAi; Notch signaling target gene; phenotype;  
KW pluripotent stem cell.  
XX  
XX Homo sapiens.  
OS  
XX WO2003068961-A2.  
PN  
XX 21-AUG-2003.  
PD  
XX 12-FEB-2003; 2003WO-GB000579.  
PF  
XX 13-FEB-2002; 2002GB-00003359.  
PR  
XX 13-FEB-2002; 2002GB-00003387.  
PR  
XX (AXOR-) AXORDIA LTD.  
PA  
XX Andrews P, Walsh J, Gokhale P;  
PI  
XX WPI; 2003-697528/66.  
XX  
XX New inhibitory RNA molecule having double stranded RNA molecules, useful  
PT for manipulating the phenotype of stem cells, preferably pluripotent  
PT stem cells.  
XX  
XX Disclosure; SEQ ID NO 128; 157pp; English.  
PS  
XX  
XX The invention relates to an inhibitory RNA (RNAi) molecule derived from a  
CC nucleic acid molecule comprising a defined nucleic acid sequences given  
CC in the specification or a sequence which hybridizes to the sequences and  
CC encodes a Notch signaling target gene or which is a degenerate as a  
CC result of the genetic code of the sequences. The methods and compositions  
CC of the present invention are useful for manipulating the phenotype of  
CC stem cells, preferably pluripotent stem cells. This sequence corresponds  
CC to one of the nucleic acid molecules of the invention.  
XX  
XX Sequence 1647 BP; 504 A; 396 C; 341 G; 406 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 0.0693 Length: 1647  
Score: 61.00 Matches: 11  
Percent Similarity: 85.71% Conservative: 1  
Best Local Similarity: 78.57% Mismatches: 2  
Query Match: 75.31% Indels: 0  
DB: 10 Gaps: 0  
AAH26175\_COPY\_50\_65 (1-16) x ADK67038 (1-1647)  
QY 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14  
DB 174 GAGATGACACTGAAATTTGGTCCAGCAGCTAATGACAGTGTC 215  
RESULT 3  
ADK66978  
ID ADK66978 standard; DNA; 1647 BP.  
XX  
AC ADK66978;  
XX  
XX 06-MAY-2004 (first entry)  
DT  
DE Gene #68 for inhibitory RNA to manipulate stem cell phenotype.  
XX  
XX ds; gene; inhibitory RNA; RNAi; Notch signaling target gene; phenotype;  
KW pluripotent stem cell.  
XX  
XX Homo sapiens.  
OS  
XX WO2003068961-A2.  
PN  
XX 21-AUG-2003.  
PD  
XX 12-FEB-2003; 2003WO-GB000579.  
PF  
XX 13-FEB-2002; 2002GB-00003359.  
PR  
XX 13-FEB-2002; 2002GB-00003387.  
PR  
XX (AXOR-) AXORDIA LTD.  
PA  
XX Andrews P, Walsh J, Gokhale P;  
PI  
XX WPI; 2003-697528/66.  
XX  
XX New inhibitory RNA molecule having double stranded RNA molecules, useful  
PT for manipulating the phenotype of stem cells, preferably pluripotent  
PT stem cells.  
XX  
XX Disclosure; SEQ ID NO 68; 157pp; English.  
PS  
XX  
XX The invention relates to an inhibitory RNA (RNAi) molecule derived from a  
CC nucleic acid molecule comprising a defined nucleic acid sequences given  
CC in the specification or a sequence which hybridizes to the sequences and  
CC encodes a Notch signaling target gene or which is a degenerate as a  
CC result of the genetic code of the sequences. The methods and compositions  
CC of the present invention are useful for manipulating the phenotype of  
CC stem cells, preferably pluripotent stem cells. This sequence corresponds  
CC to one of the nucleic acid molecules of the invention.  
XX  
XX Sequence 1647 BP; 504 A; 396 C; 341 G; 406 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 0.0693 Length: 1647  
Score: 61.00 Matches: 11  
Percent Similarity: 85.71% Conservative: 1  
Best Local Similarity: 78.57% Mismatches: 2  
Query Match: 75.31% Indels: 0  
DB: 10 Gaps: 0  
AAH26175\_COPY\_50\_65 (1-16) x ADK67038 (1-1647)  
QY 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14  
DB 174 GAGATGACACTGAAATTTGGTCCAGCAGCTAATGACAGTGTC 215  
RESULT 3  
ADK66978  
ID ADK66978 standard; DNA; 1647 BP.  
XX  
AC ADK66978;  
XX  
XX 06-MAY-2004 (first entry)  
DT  
DE Gene #68 for inhibitory RNA to manipulate stem cell phenotype.  
XX  
XX ds; gene; inhibitory RNA; RNAi; Notch signaling target gene; phenotype;  
KW pluripotent stem cell.  
XX  
XX Homo sapiens.  
OS  
XX WO2003068961-A2.  
PN  
XX 21-AUG-2003.  
PD  
XX 12-FEB-2003; 2003WO-GB000579.  
PF  
XX 13-FEB-2002; 2002GB-00003359.  
PR  
XX 13-FEB-2002; 2002GB-00003387.  
PR  
XX (AXOR-) AXORDIA LTD.  
PA  
XX Andrews P, Walsh J, Gokhale P;  
PI  
XX WPI; 2003-697528/66.  
XX  
XX New inhibitory RNA molecule having double stranded RNA molecules, useful  
PT for manipulating the phenotype of stem cells, preferably pluripotent  
PT stem cells.  
XX  
XX Disclosure; SEQ ID NO 128; 157pp; English.  
PS  
XX  
XX The invention relates to an inhibitory RNA (RNAi) molecule derived from a  
CC nucleic acid molecule comprising a defined nucleic acid sequences given  
CC in the specification or a sequence which hybridizes to the sequences and

Best Local Similarity: 78.57%  
Query Match: 75.31%  
DB: 10  
Aah26175\_COPY\_50\_65 (1-16) x ADK66978 (1-1647)  
QY 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysile 14  
DB 174 GAGATGACACTGAAATTGGTCCAGCACGTAATGACAGTGTC 215

RESULT 4  
ID ADB53173  
AD B53173 standard; DNA; 1852 BP.  
XX AC ADB53173;  
XX DT 04-DEC-2003 (first entry)  
XX DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3715.  
XX KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;  
XX KW toxicity marker; toxicity progression; drug screening;  
XX KW primary rat hepatocyte toxicity modelling; gene; ds.  
XX OS Rattus norvegicus.  
XX PN WO2003065993-A2.  
XX PD 14-AUG-2003.  
XX PF 04-FEB-2003; 2003WO-US003482.  
XX PR 04-FEB-2002; 2002US-0353171P.  
XX PR 13-MAR-2002; 2002US-0363534P.  
XX PR 08-APR-2002; 2002US-0370248P.  
XX PR 10-APR-2002; 2002US-0371134P.  
XX PR 10-APR-2002; 2002US-0371135P.  
XX PR 10-APR-2002; 2002US-0371150P.  
XX PR 11-APR-2002; 2002US-0371413P.  
XX PR 19-APR-2002; 2002US-0373601P.  
XX PR 19-APR-2002; 2002US-0373602P.  
XX PR 22-APR-2002; 2002US-0374139P.  
XX PR 08-MAY-2002; 2002US-0378370P.  
XX PR 09-MAY-2002; 2002US-0378652P.  
XX PR 09-MAY-2002; 2002US-0378653P.  
XX PR 09-MAY-2002; 2002US-0378655P.  
XX PR 09-JUL-2002; 2002US-0394230P.  
XX PR 09-JUL-2002; 2002US-0394253P.  
XX PR 04-SEP-2002; 2002US-0407688P.  
XX PR 28-JAN-2003; 2003US-0442900P.  
XX PA (GENE-) GENE LOGIC INC.

XX PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;  
XX PI Blashoff M;  
XX WPI; 2003-731472/69.  
XX DR WPI; 2003-731472/69.  
XX PT Determining if a compound induces a toxic effect on a tissue or cell, for  
XX PT identifying hepatotoxic compounds, comprises comparing a gene expression  
XX PT profile of a tissue or cell sample to a database of Tox mean and non-Tox  
XX PT mean values.

XX PS Claim 44; SEQ ID NO 3715; 874pp; English.  
XX CC The present invention describes a method for determining whether a  
XX CC compound induces a toxic effect on a tissue or cell. The method comprises  
XX CC preparing a gene expression profile of a tissue or cell sample exposed to  
XX CC the compound, and comparing the gene expression profile to a database  
XX CC comprising data or information on the Tox mean and non-Tox mean value.  
XX CC The method is useful for predicting or identifying at least one toxic  
XX CC effect, particularly hepatotoxicity, of a test or unknown compound. The  
XX CC genes listed in the specification are useful as diagnostic or toxicity

CC markers for the prediction or identification of the physiological state  
CC of tissue or cell sample that has been exposed to a compound, or to  
CC identify or predict the toxic effects of a compound or an agent. These  
CC may also be used as markers for monitoring toxicity progression or for  
CC drug screening. The present sequence represents a primary rat hepatocyte  
CC toxicity modelling related gene sequence from the present invention.  
XX Sequence 1852 BP; 558 A; 472 C; 405 G; 417 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 0.08 Length: 1852  
Score: 61.00 Matches: 11  
Percent Similarity: 85.71% Conservative: 1  
Best Local Similarity: 78.57% Mismatches: 2  
Query Match: 75.31% Indels: 0  
DB: 10 Gaps: 0

Aah26175\_COPY\_50\_65 (1-16) x ADB53173 (1-1852)  
QY 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysile 14  
DB 432 GAGATGACACTGAAATTGGTCCAGCACGTAATGACAGTGTC 473

RESULT 5  
ID ADC22243  
AC ADC22243 standard; DNA; 51 BP.  
XX AC ADC22243;  
XX DT 18-DEC-2003 (first entry)  
XX DE Protein binding domain nucleotide sequence SEQ ID NO:92.  
XX KW recombinant fusion protein; fusion protein; binding; detection;  
XX KW localisation domain; binding domain;  
XX KW subcellular compartment localisation; gene; ds.  
XX OS Homo sapiens.  
XX PN WO2003012068-A2.  
XX PD 13-FEB-2003.  
XX PF 01-AUG-2002; 2002WO-US024572.  
XX PR 01-AUG-2001; 2001US-0309395P.  
XX PR 13-DEC-2001; 2001US-0341589P.  
XX PA (CELL-) CELLOMICS INC.

XX PI Bright G, Premkumar DR, Chen Y;  
XX WPI; 2003-248174/24.  
XX DR P-PSDB; ADC22242.  
XX PT New recombinant fusion protein comprising detection and first  
XX PT localisation domains and a binding domain for the molecule of interest,  
XX PT useful for detecting binding of a molecule of interest.  
XX PS Disclosure; SEQ ID NO 92; 101pp; English.

XX CC The present invention describes a recombinant fusion protein (I) for  
XX CC detecting binding of a molecule of interest. (I) comprises: (a) a  
XX CC detection domain; (b) a first localisation domain; and (c) a binding  
XX CC domain for the molecule of interest. The detection domain, the first  
XX CC localisation domain and the binding domain for the molecule of interest  
XX CC constituting the recombinant fusion protein for detecting binding of a  
XX CC molecule of interest are operably linked. The binding domain for the  
XX CC molecule of interest is separated from the first localisation domain by 0  
XX CC -20 amino acid residues. The first localisation domain and the binding  
XX CC domain for the molecule of interest both do not occur in a single non-  
XX CC recombinant protein with the same spacing as in the recombinant fusion  
XX CC protein for detecting binding of a molecule of interest. Also described:

CC (1) a recombinant nucleic acid encoding the recombinant fusion protein;  
 CC (2) a recombinant expression vector comprising the nucleic acid control  
 CC sequences operably linked to the recombinant nucleic acid molecule; (3) a  
 CC genetically engineered host cell transfected with the recombinant  
 CC expression vector; (4) a kit for detecting binding of the molecule of  
 CC interest; and (5) a method for identifying compounds that alter the  
 CC binding of the molecule of interest. The recombinant fusion protein is  
 CC useful for detecting binding of a molecule of interest. The recombinant  
 CC fusion protein eliminates the need to construct two or more chimeric  
 CC proteins and enables the monitoring of biochemical events in live, intact  
 CC or fixed cells. The present sequence is used in the exemplification of  
 CC the present invention.

SQ Sequence 51 BP; 20 A; 10 C; 9 G; 12 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 0.00398 Length: 51  
 Score: 58.00 Matches: 11  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 71.60% Indels: 0  
 DB: 10 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x ADC22243 (1-51)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAan 11  
 DB 19 GAGATGACACTGAAATTTGGTCCAGCAGTAAT 51

RESULT 6

AAH26175  
 ID AAS61674 standard; cDNA; 501 BP.

XX AAS61674;

AC AAS61674;

DT 29-JAN-2002 (first entry)

XX Lung small cell carcinoma antigen, cDNA #215.

KW Human; cytostatic; antitumour; lung small cell cancer antigen; tumour;

KW lung cancer; ss.

XX Homo sapiens.

XX WO200177168-A2.

XX 18-OCT-2001.

XX 11-APR-2001; 2001WO-US011859.

XX 11-APR-2000; 2000US-0196780P.

XX 21-JUN-2000; 2000US-0213361P.

XX 01-SEP-2000; 2000US-0229763P.

XX 05-SEP-2000; 2000US-0230629P.

XX 14-SEP-2000; 2000US-0232565P.

XX 19-DEC-2000; 2000US-0257037P.

XX 08-JAN-2001; 2001US-0260796P.

XX (CORI-) CORIXA CORP.

XX Lodes MJ, Wang T, Mohamath R, Indirias CY;

XX WPI; 2002-010896/01.

XX Lung tumor polynucleotide and polypeptides useful in therapy and

XX diagnosis of cancer especially lung cancer.

XX Claim 1; Page 195; 295pp; English.

CC contacted with (III), detecting the amount of polynucleotide hybridised  
 CC to (III) in the sample and comparing the amount of polynucleotide to a  
 CC predetermined cut-off value and thereby determining cancer in a patient.  
 CC (I), (II) or antigen-presenting cells expressing (II) is useful for  
 CC stimulating and/or expanding T cells with one of the components under  
 CC method comprises contacting T cells with one of the components under  
 CC conditions to permit the stimulation and/or expansion of the cells. A  
 CC composition comprising (I) is useful for stimulating an immune response  
 CC in a patient and for inhibiting the development of a cancer especially  
 CC lung cancer in a patient. An isolated T cell population is useful for  
 CC removing tumour cells from the biological sample and for inhibiting the  
 CC development of cancer in a patient. AAS61460-AAS61874 represent novel  
 CC human lung small cell cancer antigen coding sequences of the invention

SQ Sequence 501 BP; 151 A; 103 C; 121 G; 125 T; 0 U; 1 Other;

Alignment Scores:  
 Pred. No.: 0.102 Length: 501  
 Score: 57.00 Matches: 10  
 Percent Similarity: 78.57% Conservative: 1  
 Best Local Similarity: 71.43% Mismatches: 3  
 Query Match: 70.37% Indels: 0  
 DB: 6 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x AAS61674 (1-501)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAanSerLysile 14  
 DB 414 GAGATGACACTGAAATTTGGTCCANCAAGTATGACAGTGC 455

RESULT 7

AAH26175

ID AAC02948 standard; cDNA; 229 BP.

XX AAC02948;

AC AAC02948;

DT 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 2946.

DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-00200610.

XX 26-FEB-1999; 99US-0122487P.

XX (GEST ) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX P-PSDB; AAG02942.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for

XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for

XX diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX Claim 1; SEQ ID NO 2946; 71pp + Sequence Listing; English.

XX The present sequence is one of a large number of 5' ESTs derived from

XX mRNAs encoding secreted proteins. An ORF has been identified within the

XX sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs

XX derived from 30 different tissues. EST sequences usually correspond

XX mainly to the 3' untranslated region (UTR) of the mRNA because they are

XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not

XX well suited for isolating cDNA sequences derived from the 5' ends of

CC mRNAs and even in those cases where longer cDNA sequences have been  
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
CC mRNAs with intact 5' ends and can therefore be used to obtain full length  
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
CC gene therapy and chromosome mapping procedures. They are used to obtain  
CC upstream regulatory sequences and to design expression and secretion  
CC vectors

SQ Sequence 229 BP; 63 A; 50 C; 67 G; 48 T; 0 U; 1 Other;

#### Alignment Scores:

Pred. No.: 0.0989 Length: 229  
Score: 55.00 Matches: 10  
Percent Similarity: 78.57% Conservative: 1  
Best Local Similarity: 71.43% Mismatches: 3  
Query Match: 67.90% Indels: 0  
DB: 3 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x AAC02948 (1-229)

QY 1 GluMetThrLeuIysPheGlyProAlaArgAsnSerLysIle 14  
|||||  
DB 165 GAGATGACATTGAAATTTGGCCCGCCGAACTGACTCAGTC 206  
:::

#### RESULT 8

ABV28993  
ID ABV28993 standard; cDNA; 800 BP.

XX AC ABV28993;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 28984.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PS Claim 1; Page 6132; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX SQ Sequence 800 BP; 244 A; 147 C; 174 G; 224 T; 0 U; 11 Other;

#### Alignment Scores:

Pred. No.: 0.455 Length: 800  
Score: 55.00 Matches: 10  
Percent Similarity: 78.57% Conservative: 1  
Best Local Similarity: 71.43% Mismatches: 3  
Query Match: 67.90% Indels: 0  
DB: 5 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x ABV28993 (1-800)

QY 1 GluMetThrLeuIysPheGlyProAlaArgAsnSerLysIle 14  
|||||  
DB 179 GAGATGACATTGAAATTTGGCCCGCCGAACTGACTCAGTC 220  
:::

#### RESULT 9

ABV23152  
ID ABV23152 standard; cDNA; 800 BP.

XX AC ABV23152;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 23143.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PS Claim 1; Page 4152; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX

SQ Sequence 800 BP; 244 A; 147 C; 174 G; 224 T; 0 U; 11 Other;

Alignment Scores:  
 Pred. No.: 0.455 Length: 800  
 Score: 55.00 Matches: 10  
 Percent Similarity: 78.57% Conservative: 1  
 Best Local Similarity: 71.43% Mismatches: 3  
 Query Match: 67.90% Indels: 0  
 DB: 5 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x ABV23152 (1-800)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAanSerLysIle 14  
 |||||  
 DB 179 GAGATGACACTGAAATTTGGCCCGACGCGACTGACTCAGTC 220  
 :::

RESULT 10  
 ADP80924  
 ID ADP80924 standard; cDNA; 1239 BP.

XX AC ADP80924;  
 XX DT 09-SEP-2004 (first entry)  
 XX DE Mouse C130020M04Rik nucleotide sequence SEQ ID NO:149.

XX KW c-fos; c-fos interacting protein; fos interacting protein chromosome X;  
 XX KW Fip-cx; screening; Fc-fos interacting protein inhibitor; mouse; gene; ss.  
 XX OS Mus musculus.

XX FH Key Location/Qualifiers  
 XX FT CDS 1..1239  
 XX FT /\*tag= a  
 XX FT /product= "C130020M04Rik"

XX PN WO2004053121-A1.

XX PD 24-JUN-2004.

XX PF 19-NOV-2003; 2003WO-JP014749.

XX PR 11-DEC-2002; 2002JP-00360046.

XX PA (UYKE-) UNIV KEIO.

XX PI Miyamoto E, Ishizaka M, Yanagawa H;

XX DR WPI; 2004-517250/49.

XX DR P-PSDB; ADP80860.

XX PT New proteins that interact with fos, e.g., fos interacting protein  
 XX PT chromosome X (Fip-cx).

XX PS Claim 59; SEQ ID NO 149; 192pp; Japanese.

XX CC The present invention describes a protein (I) that interacts with c-fos  
 CC (e.g., fos interacting protein chromosome X (Fip-cx), Fip-cx.1, fos  
 CC interacting protein chromosome ex.2, or fos interacting protein  
 CC chromosome 4). Also described: (1) a nucleic acid (II) that encodes (I);  
 CC (2) an inhibitor (III) that inhibits the interaction of (I) encoded by  
 CC (II) with c-fos protein; (3) an inhibitor (IV) that inhibits interaction  
 CC of a protein with c-fos protein; and (4) detecting (M1) the interaction  
 CC of a protein as mentioned in (3) with c-fos protein, by contacting the  
 CC protein with c-fos, to form a composite. (I) is useful for detecting its  
 CC interaction with c-fos, to form a complex. (M1) is useful for screening  
 CC the protein that interacts with c-fos which involves performing the  
 CC detection process and selecting the protein that interacts with c-fos.  
 CC (M1) is useful for screening a protein that interacts with c-fos which  
 CC involves performing (M1) and selecting the detected protein. (I) is  
 CC useful for screening inhibitors that interact with c-fos. The present  
 CC sequence encodes a mouse C130020M04Rik amino acid sequence, which can  
 CC interact with c-fos in the exemplification of the present invention.

XX SQ Sequence 1239 BP; 327 A; 359 C; 282 G; 271 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 0.775 Length: 1239  
 Score: 55.00 Matches: 10  
 Percent Similarity: 78.57% Conservative: 1  
 Best Local Similarity: 71.43% Mismatches: 3  
 Query Match: 67.90% Indels: 0  
 DB: 12 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x ADP80924 (1-1239)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAanSerLysIle 14  
 |||||  
 DB 94 GAGATGACACTGAAATTTGGCCCGACGCGACTCAGTC 135  
 :::

RESULT 11  
 ABV28463  
 ID ABV28463 standard; cDNA; 2755 BP.

XX AC ABV28463;  
 XX DT 16-SEP-2002 (first entry)  
 XX DE Human prostate expression marker cDNA 28454.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 XX KW Pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of  
 XX PT prostate cells and correlating with presence of prostate cancer, useful  
 XX PT for detecting presence of prostate cancer, stage of prostate cancer.

XX PS Claim 1; Page 5940-5941; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX SQ Sequence 2755 BP; 645 A; 811 C; 612 G; 684 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.: 2.06 Length: 2755  
 Score: 55.00 Matches: 10  
 Percent Similarity: 78.57% Conservative: 1  
 Best Local Similarity: 71.43% Mismatches: 3  
 Query Match: 67.90% Indels: 0  
 DB: 5 Gaps: 0

Aah26175\_COPY\_50\_65 (1-16) x ABV28463 (1-2755)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAenSerLysIle 14  
 |||||  
 DB 194 GAGATGACATTGAAATTTGGCCAGCCGAACTGACTCAGTC 235  
 ::::

RESULT 12  
 ABV22640  
 ID ABV22640 standard; cDNA; 2755 BP.  
 XX  
 AC ABV22640;  
 XX  
 DT 13-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker cDNA 22631.  
 XX  
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US005171.  
 XX  
 PR 17-FEB-2000; 2000US-0183319P.  
 PR 16-MAR-2000; 2000US-0189862P.  
 PR 25-MAY-2000; 2000US-0207454P.  
 PR 09-JUN-2000; 2000US-0211314P.  
 PR 18-JUL-2000; 2000US-0219007P.  
 PR 13-DEC-2000; 2000US-0255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 PI Schlegel R, Endege WO, Monahan JE;  
 DR WPI; 2001-662795/76.  
 XX  
 PT Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.  
 XX  
 PS Claim 1; Page 3967-3968; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the  
 CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX  
 SQ Sequence 2755 BP; 645 A; 811 C; 612 G; 684 T; 0 U; 3 Other;

Alignment Scores:  
 Pred. No.: 2.06 Length: 2755  
 Score: 55.00 Matches: 10  
 Percent Similarity: 78.57% Conservative: 1  
 Best Local Similarity: 71.43% Mismatches: 3

Query Match: 67.90% Indels: 0  
 DB: 5 Gaps: 0

Aah26175\_COPY\_50\_65 (1-16) x ABV22640 (1-2755)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAenSerLysIle 14  
 |||||  
 DB 194 GAGATGACATTGAAATTTGGCCAGCCGAACTGACTCAGTC 235  
 ::::

RESULT 13  
 ACF87477  
 ID ACF87477 standard; DNA; 2758 BP.  
 XX  
 AC ACF87477;  
 XX  
 DT 02-JUN-2005 (first entry)  
 XX  
 DE Human SIRS/sepsis diagnostic marker DNA fragment 6337.  
 XX  
 KW Systemic inflammatory response syndrome; SIRS; antibacterial;  
 KW immunosuppressive; antiinflammatory; diagnosis; sepsis; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004087949-A2.  
 XX  
 PD 14-OCT-2004.  
 XX  
 PF 31-MAR-2004; 2004WO-EP003419.  
 XX  
 PR 02-APR-2003; 2003DB-01015031.  
 PR 08-AUG-2003; 2003DE-01036511.  
 PR 02-SEP-2003; 2003DE-01040395.  
 XX  
 PA (SIRS-) SIRS LAB GMBH.  
 XX  
 PI Russwurm S, Reinhart K, Saluz H, Straube E, Zipfel PF, Deigner H;  
 XX WPI; 2004-748070/73.  
 DR  
 XX  
 PT In vitro detection of systemic inflammatory response syndrome and related  
 PT conditions, for e.g. monitoring progression, comprises detecting abnormal  
 PT expression of disease-related genes.  
 XX  
 PS Disclosure; Page; 75pp; German.  
 XX  
 CC The invention relates to a novel method for in vitro detection of  
 CC systemic inflammatory response syndrome (SIRS). The method comprises  
 CC detecting abnormal expression of disease-related genes, or their  
 CC associated peptides. The method of the invention demonstrates  
 CC antibacterial, immunosuppressive and antiinflammatory applications and  
 CC may be used for early differential diagnosis, monitoring progression,  
 CC assessing risk, assessing the likely response to treatment and for post  
 CC mortem diagnosis of systemic inflammatory response syndrome, sepsis and  
 CC sepsis-like conditions. The recombinant or synthetic nucleic acid  
 CC sequences of the invention, or derived proteins or peptides, may be  
 CC useful as calibrants in assays for the specified diseases, for evaluating  
 CC activity or toxicity in screening for active agents and/or for  
 CC preparation of agents for treatment or prevention of the specified  
 CC diseases. The current sequence is that of a human SIRS/sepsis diagnostic  
 CC marker DNA fragment of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at ftp.wipo.int/pub/published  
 CC pct\_sequences. Furthermore, a number of arbitrary SEQ ID NO.8 are  
 CC disclosed within the specification, however, these have not been taken  
 CC into account during indexing due to inconsistencies in application and  
 CC format  
 XX  
 SQ Sequence 2758 BP; 647 A; 813 C; 612 G; 686 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2.06 Length: 2758  
 Score: 55.00 Matches: 10

Percent Similarity: 78.57%  
Best Local Similarity: 71.43%  
Query Match: 13  
DB: 13  
Conservative: 1  
Mismatch: 3  
Indels: 0  
Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x ACF87477 (1-2758)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysLeu 14  
DB 197 GAGATGACATTGAAATTTGGCCCGCCGACTGACTCAGTC 238

RESULT 14  
ADC22231  
ID ADC22231 standard; DNA; 51 BP.  
AC ADC22231;  
DT 18-DEC-2003 (first entry)  
XX Protein binding domain nucleotide sequence SEQ ID NO:80.  
DE recombinant fusion protein; fusion protein; binding; detection;  
KW localisation domain; binding domain;  
KW subcellular compartment localisation; gene; ds.  
OS Homo sapiens.  
XX WO2003012068-A2.  
PN 13-FEB-2003.  
XX 01-AUG-2002; 2002WO-US024572.  
XX 01-AUG-2001; 2001US-0309395P.  
PR 13-DEC-2001; 2001US-0341589P.  
XX (CELL-) CELLOMICS INC.  
PA Bright G, Premkumar DR, Chen Y;  
PI WPI; 2003-248174/24.  
DR P-PSDB; ADC22230.  
XX New recombinant fusion protein comprising detection and first  
PT localization domains and a binding domain for the molecule of interest,  
PT useful for detecting binding of a molecule of interest.  
XX Disclosure; SEQ ID NO 80; 101pp; English.

The present invention describes a recombinant fusion protein (I) for  
detecting binding of a molecule of interest. (I) comprises: (a) a  
detection domain; (b) a first localisation domain; and (c) a binding  
domain for the molecule of interest. The detection domain, the first  
localisation domain and the binding domain for the molecule of interest  
constituting the recombinant fusion protein for detecting binding of a  
molecule of interest are operably linked. The binding domain for the  
molecule of interest is separated from the first localisation domain by 0  
-20 amino acid residues. The first localisation domain and the binding  
domain for the molecule of interest both do not occur in a single non-  
recombinant protein with the same spacing as in the recombinant fusion  
protein for detecting binding of a molecule of interest. Also described:  
(1) a recombinant nucleic acid encoding the recombinant fusion protein;  
(2) a recombinant expression vector comprising the nucleic acid control;  
(3) sequences operably linked to the recombinant nucleic acid molecule;  
(4) a genetically engineered host cell transfected with the recombinant  
expression vector; (5) a method for identifying compounds that alter the  
binding of the molecule of interest. The recombinant fusion protein is  
useful for detecting binding of a molecule of interest. The recombinant  
fusion protein eliminates the need to construct two or more chimeric  
proteins and enables the monitoring of biochemical events in live, intact  
or fixed cells. The present sequence is used in the exemplification of  
the present invention.

Alignment Scores:  
Sequence 397 BP; 118 A; 87 C; 96 G; 96 T; 0 U; 0 Other;

Percent Similarity: 78.57%  
Best Local Similarity: 71.43%  
Query Match: 13  
DB: 13  
Conservative: 1  
Mismatch: 3  
Indels: 0  
Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x ADC22231 (1-51)

QY 1 GluMetThrLeuLysPheGlyProAlaArg 10  
DB 19 GAGATGACATTGAAATTTGGCCCGCCGACTGACTCAGTC 48

RESULT 15  
ABV14146  
ID ABV14146 standard; cDNA; 397 BP.  
XX AC ABV14146;  
XX DT 13-SEP-2002 (first entry)  
XX Human prostate expression marker cDNA 14137.  
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX Homo sapiens.  
XX WO200160860-A2.  
PN 23-AUG-2001.  
XX 20-FEB-2001; 2001WO-US005171.  
XX 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PA Schlegel R, Endege WO, Monahan JE;  
PI WPI; 2001-662795/76.  
DR Novel isolated nucleic acid molecule associated with cancerous state of  
CC prostate cells and correlating with presence of prostate cancer, useful  
CC for detecting presence of prostate cancer, stage of prostate cancer.  
XX Claim 1; Page 2358; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising  
a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
specification or its complement. (I) is useful for: (a) assessing whether  
a patient is afflicted with prostate cancer; (b) monitoring the efficacy  
of a test compound to inhibit prostate cancer in a patient; (c) assessing the  
progression of prostate cancer in a patient; (d) assessing the efficacy  
of a test compound to inhibit prostate cancer in a patient; (e) assessing  
the efficacy of a therapy for inhibiting prostate cancer in a patient;  
(f) selecting a composition for inhibiting prostate cancer in a patient;  
(g) assessing the prostate cell carcinogenic potential of a compound; (h)  
determining whether prostate cancer has metastasized in a patient; (h)  
assessing the aggressiveness or indolence of prostate cancer in a patient  
; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

Alignment Scores:  
Sequence 397 BP; 118 A; 87 C; 96 G; 96 T; 0 U; 0 Other;



Pred. No.: 12.2 Length: 397  
 Score: 46.00 Matches: 9  
 Percent Similarity: 90.00% Conservative: 0  
 Best Local Similarity: 90.00% Mismatches: 1  
 Query Match: 56.79% Indels: 0  
 DB: 5 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x ABV14146 (1-397)

Qy 1 GluMetThrLeuLysPheGlyProAlaArg 10  
 ||||||||||||||||||||  
 Db 162 GAGATGACATTGAATTTGCCCATACCGA 191

Search completed: January 12, 2006, 19:11:14  
 Job time : 406 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2006, 17:58:06 ; Search.time 2606 Seconds  
(without alignments)

287.258 Million cell updates/sec

Title: AAh26175\_COPY\_50\_65

Perfect score: 81

Sequence: 1 emtlkfgparnsklee 16

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+pn.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/RAWLINGS10076905/runat\_10012006\_144824\_5333/app\_query.fasta\_1.199  
-DB=EST -QFMT=FASTAP -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=oct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=RAWLINGS10076905 @CGN\_1\_1\_5315 @runat\_10012006\_144824\_5333 -NCPU=6  
-ICPU=3 -NO\_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*

1: gb\_est1:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	75.3	340	AA414544	AA414544 vc60a12.s
2	61	75.3	344	10	CL569431 AM0785 Sa
3	61	75.3	417	9	CC249614 XI426 Bay
4	61	75.3	506	1	AJ448042 AJ448042
5	61	75.3	508	7	CR542857 DKFZP459L
6	61	75.3	515	6	CD693411 EST9934 h
7	61	75.3	538	1	AU279581 AU279581

8	61	75.3	571	1	AJ450203	AJ450203 AJ450203
9	61	75.3	576	3	BP257160	BP257160 BP257160
10	61	75.3	577	7	CN367889	CN367889 170004241
11	61	75.3	582	3	BP201994	BP201994 BP201994
12	61	75.3	606	3	BQ037883	BQ037883 pgnic.pk0
13	61	75.3	613	5	BU662818	BU662818 C191901.2
14	61	75.3	621	7	CN367892	CN367892 170005326
15	61	75.3	635	1	AL119505	AL119505 DKFZP7610
16	61	75.3	637	5	BY736938	BY736938 BY736938
17	61	75.3	640	7	CN455108	CN455108 UI-M-HN0-
18	61	75.3	642	7	CR628862	CR628862 DKFZP469C
19	61	75.3	650	7	CN367893	CN367893 170005326
20	61	75.3	651	6	CB438334	CB438334 686887 MA
21	61	75.3	651	7	CN367890	CN367890 170005319
22	61	75.3	653	2	BB627898	BB627898 BB627898
23	61	75.3	662	6	CD355777	CD355777 UI-M-FY0-
24	61	75.3	664	6	CF535749	CF535749 UI-M-GH0-
25	61	75.3	669	2	BB665053	BB665053 BB665053
26	61	75.3	669	5	BY734700	BY734700 BY734700
27	61	75.3	669	5	BY739920	BY739920 BY739920
28	61	75.3	670	7	CN367891	CN367891 170004247
29	61	75.3	686	2	BE882503	BE882503 601507056
30	61	75.3	716	2	BF613172	BF613172 DB07C03.Y
31	61	75.3	722	6	CF532094	CF532094 UI-M-FY0-
32	61	75.3	723	7	CN454827	CN454827 UI-M-HN0-
33	61	75.3	737	8	CX239019	CX239019 NMA06762
34	61	75.3	742	7	CK634629	CK634629 UI-M-HN0-
35	61	75.3	744	1	AJ453735	AJ453735 AJ453735
36	61	75.3	746	6	CD348287	CD348287 UI-M-FY0-
37	61	75.3	749	6	CD348288	CD348288 UI-M-FY0-
38	61	75.3	758	6	CA510974	CA510974 UI-R-FJ0-
39	61	75.3	758	6	CB521135	CB521135 UI-M-GH0-
40	61	75.3	781	6	CA512317	CA512317 UI-R-FJ0-
41	61	75.3	782	6	CD348295	CD348295 UI-M-FY0-
42	61	75.3	792	6	CA750476	CA750476 UI-M-FY0-
43	61	75.3	802	7	CR629564	CR629564 DKFZP469G
44	61	75.3	805	1	AJ454962	AJ454962 AJ454962
45	61	75.3	806	6	CB961368	CB961368 AGENCOURT

#### ALIGNMENTS

AA414544 340 bp mRNA linear EST 04-AUG-1997  
vc60a12.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone  
IMAGE:778942 5', similar to gb:M7167 Mouse T-cell antigen receptor  
alpha-chain (MOUSE) ;, mRNA sequence.

AA414544

AA414544.1 GI:2074680

EST.

Mus musculus (house mouse)

Mus musculus

Mus musculus

EST.

1 (bases 1 to 340)

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

The WashU-HM Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HM Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

```

MG1:471798
Seq primer: -40ml3 fwd, ET from Amersham.
Location/Qualifiers
1. .340
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:778942"
/tissue_type="embryo"
/dev_stage="2-cell"
/lab_host="DH10B"
/clone_lib="Knowles Solter mouse 2 cell"
/note="Organ: embryo; Vector: pbluescribe (modified);
Site_1: MluI; Site_2: SalI; Cloned unidirectionally from
mRNA prepared from 13,500 2-cell stage embryos. Primer:
SalI (dT): 5'-CGTGCACGTCGACGTTTTTTTTTTT-3'. cDNAs
were cloned into the MluI/SalI sites of a modified
pbluescribe vector using commercial linkers (NEB).
Average insert size: 1.2 kb."
Alignment Scores:
Pred. No.: 0.199 Length: 340
Score: 61.00 Matches: 11
Percent Similarity: 85.71% Conservative: 1
Best Local Similarity: 78.57% Mismatches: 2
Query Match: 75.31% Indels: 0
DB: 1 Gaps: 0
AAH26175_COPY_50_65 (1-16) x AA414544 (1-340)
ORIGIN
1 GluMetThrLeuLysPheGlyProAlaArgAenSerLysIle 14
|||||
Db 79 GAGATGACACTGAAATTTGGTCCAGCAGTAATGACAGTGTC 120

RESULT 2
LOCUS CL569431
DEFINITION CL569431.5 GI:62469378
ACCESSION CL569431
VERSION CL569431.5
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
CL569431 344 bp mRNA linear GSS 11-APR-2005
AM0785 Sanger Institute Gene Trap Library pGT01xr Mus musculus
cDNA, mRNA sequence.
CL569431
CL569431.5 GI:62469378
GSS.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 344)
Sanger Institute Gene Trap Resource - SIGTR.
http://www.sanger.ac.uk/PostGenomics/genetrap/
Unpublished (2003)
On Apr 11, 2005 this sequence version replaced gi:60284136.
Contact: Sanger Institute Gene Trap Resource - SIGTR
Wellcome Trust Sanger Institute
Email: info.genetraps@sanger.ac.uk
Sequence tag generated by 5' RACE of total RNA from gene trap ES
cell line. ES cell lines harboring insertion mutation of target
gene are available upon request from Sanger Institute Gene Trap
Resource. Annotation information available from
http://www.sanger.ac.uk/PostGenomics/genetrap/
Class: Gene Trap.
Location/Qualifiers
1. .344
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129 OLA"
/db_xref="taxon:10090"
/sex="Male"
/cell_type="Embryonic Stem Cell"
/clone_lib="Sanger Institute Gene Trap Library pGT01xr"
/note="Vector: pGT01xr"
FEATURES
source
ORIGIN
1 GluMetThrLeuLysPheGlyProAlaArgAenSerLysIle 14
|||||
Db 79 GAGATGACACTGAAATTTGGTCCAGCAGTAATGACAGTGTC 120

RESULT 2
LOCUS CL569431
DEFINITION CL569431.5 GI:62469378
ACCESSION CL569431
VERSION CL569431.5
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
CL569431 344 bp mRNA linear GSS 11-APR-2005
AM0785 Sanger Institute Gene Trap Library pGT01xr Mus musculus
cDNA, mRNA sequence.
CL569431
CL569431.5 GI:62469378
GSS.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 344)
Sanger Institute Gene Trap Resource - SIGTR.
http://www.sanger.ac.uk/PostGenomics/genetrap/
Unpublished (2003)
On Apr 11, 2005 this sequence version replaced gi:60284136.
Contact: Sanger Institute Gene Trap Resource - SIGTR
Wellcome Trust Sanger Institute
Email: info.genetraps@sanger.ac.uk
Sequence tag generated by 5' RACE of total RNA from gene trap ES
cell line. ES cell lines harboring insertion mutation of target
gene are available upon request from Sanger Institute Gene Trap
Resource. Annotation information available from
http://www.sanger.ac.uk/PostGenomics/genetrap/
Class: Gene Trap.
Location/Qualifiers
1. .344
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129 OLA"
/db_xref="taxon:10090"
/sex="Male"
/cell_type="Embryonic Stem Cell"
/clone_lib="Sanger Institute Gene Trap Library pGT01xr"
/note="Vector: pGT01xr"
FEATURES
source
ORIGIN
1 GluMetThrLeuLysPheGlyProAlaArgAenSerLysIle 14
|||||
Db 79 GAGATGACACTGAAATTTGGTCCAGCAGTAATGACAGTGTC 120

MG1:471798
Seq primer: -40ml3 fwd, ET from Amersham.
Location/Qualifiers
1. .340
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:778942"
/tissue_type="embryo"
/dev_stage="2-cell"
/lab_host="DH10B"
/clone_lib="Knowles Solter mouse 2 cell"
/note="Organ: embryo; Vector: pbluescribe (modified);
Site_1: MluI; Site_2: SalI; Cloned unidirectionally from
mRNA prepared from 13,500 2-cell stage embryos. Primer:
SalI (dT): 5'-CGTGCACGTCGACGTTTTTTTTTTT-3'. cDNAs
were cloned into the MluI/SalI sites of a modified
pbluescribe vector using commercial linkers (NEB).
Average insert size: 1.2 kb."
Alignment Scores:
Pred. No.: 0.202 Length: 344
Score: 61.00 Matches: 11
Percent Similarity: 85.71% Conservative: 1
Best Local Similarity: 78.57% Mismatches: 2
Query Match: 75.31% Indels: 0
DB: 1 Gaps: 0
AAH26175_COPY_50_65 (1-16) x CL569431 (1-344)
ORIGIN
1 GluMetThrLeuLysPheGlyProAlaArgAenSerLysIle 14
|||||
Db 45 GAGATGACACTGAAATTTGGTCCAGCAGTAATGACAGTGTC 86

RESULT 3
LOCUS CC249614
DEFINITION CC249614.2 GI:46014514
ACCESSION CC249614
VERSION CC249614.2
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
CC249614 417 bp mRNA linear GSS 01-APR-2004
XI426 BayGenomics Gene Trap Library pGT1cd72 Mus musculus cDNA,
mRNA sequence.
CC249614
CC249614.2 GI:46014514
GSS.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 417)
BayGenomics.
http://baygenomics.ucsf.edu/
Unpublished (2001)
On Apr 1, 2004 this sequence version replaced gi:30586364.
Contact: BayGenomics
Bay Area Functional Genomics Consortium (BayGenomics)
Email: info@baygenomics.ucsf.edu
Sequence tag generated by 5' RACE of total RNA from gene trap ES
cell line. ES cell lines harboring insertion mutation of target
gene are available upon request from BayGenomics. Annotation
information available from
http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?
OPTION=EXACT&TYPE=CELL_LINE&KEY=XI426
Class: Gene Trap.
Location/Qualifiers
1. 417
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129 Ola"
/db_xref="taxon:10090"
/sex="Male"
/cell_type="Embryonic stem cell"
/clone_lib="BayGenomics Gene Trap Library pGT1cd72"
/note="Vector: pGT1cd72"
FEATURES
source
ORIGIN
1 GluMetThrLeuLysPheGlyProAlaArgAenSerLysIle 14
|||||
Db 120 GAGATGACACTGAAATTTGGTCCAGCAGTAATGACAGTGTC 161

RESULT 4
LOCUS AJ448042
DEFINITION AJ448042.1 GI:100000000
ACCESSION AJ448042
VERSION AJ448042.1
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
AJ448042 506 bp mRNA linear EST 19-APR-2002

```

```

DEFINITION AJ448042 riken1 Gallus gallus cdna clone 18113r1, mRNA sequence.
ACCESSION AJ448042
VERSION AJ448042.1 GI:20215263
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 506)
Buerstedde,J.M.
AUTHORS Gallus gallus bursal lymphocyte EST
TITLE Gallus gallus bursal lymphocyte EST
JOURNAL Unpublished (2002)
COMMENT Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.

FEATURES
source
1..506
Location/Qualifiers
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="18113r1"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/clone_lib="riken1"
/note="TCB inbred strain"

ORIGIN
Alignment Scores:
Pred. No.: 0.313 Length: 506
Score: 61.00 Matches: 11
Percent Similarity: 85.71% Conservat: 1
Best Local Similarity: 78.57% Mismatches: 2
Query Match: 75.31% Indels: 0
DB: 1 Gaps: 0

AAH26175_COPY_50_65 (1-16) x AJ448042 (1-506)
Qy 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14
Db 433 GAGATGACACTGAATTTGGTCCGCGCTGATGATGATGTC 474
|||||
|:::

RESULT 5
CR542857 508 bp mRNA linear EST 07-JUL-2004
LOCUS DKFZp459L2240_r1 459 (synonym: pcorl) Pongo pygmaeus cdna clone
DEFINITION DKFZp459L2240 5', mRNA sequence.
ACCESSION CR542857
VERSION CR542857.1 GI:49894012
KEYWORDS EST.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Pongo.
1 (bases 1 to 508)
Bloeker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B., Amid,C.,
Osanger,A., Fobo,G., Han,M. and Wiemann,S.
Pongo pygmaeus mRNA (Bloeker,H., Boecher,M., Brandt,P., et al.)
Unpublished (2004)
CONTACT: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert from S. Wiemann,
Molecular Genome Analysis, German Cancer Research Center (DKFZ);
Email a.wiemann@dkfz-heidelberg.de; sequenced by GSF (National
Research Centre for Biotechnology Ltd., Braunschweig/Germany)
within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp459L2240) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further

```

```

information about the clone and the sequencing project is available
at http://mips.gsf.de/projects/cdna/.

FEATURES
source
1..508
Location/Qualifiers
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp459L2240"
/tissue_type="cortex"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="459 (synonym: pcorl)"
/note="Vector: pSport1_Sfi; Site_1: SfiIa; Site_2: SfiIb"

ORIGIN
Alignment Scores:
Pred. No.: 0.314 Length: 508
Score: 61.00 Matches: 11
Percent Similarity: 85.71% Conservat: 1
Best Local Similarity: 78.57% Mismatches: 2
Query Match: 75.31% Indels: 0
DB: 7 Gaps: 0

AAH26175_COPY_50_65 (1-16) x CR542857 (1-508)
Qy 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14
Db 459 GAGATGACACTGAATTTGGTCCGCGCTGATGATGATGTC 500
|||||
|:::

RESULT 6
CD693411 515 bp mRNA linear EST 25-JUN-2003
LOCUS EST9934 human nasopharynx Homo sapiens cdna, mRNA sequence.
DEFINITION CD693411
ACCESSION CD693411
VERSION CD693411.1 GI:32217044
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
1 (bases 1 to 515)
Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
Zeng,Y.-X.
Transcriptional Gene Expression Profile of Human Nasopharynx
Unpublished (2003)
CONTACT: Yixin Zeng
Cancer Center
Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China
Tel: 86-1380-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@gzsunm.edu.cn.

FEATURES
source
1..515
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cdna
library from southern Chinese"

ORIGIN
Alignment Scores:
Pred. No.: 0.319 Length: 515
Score: 61.00 Matches: 11
Percent Similarity: 85.71% Conservat: 1
Best Local Similarity: 78.57% Mismatches: 2
Query Match: 75.31% Indels: 0
DB: 6 Gaps: 0

AAH26175_COPY_50_65 (1-16) x CD693411 (1-515)

```

```

QY      1  GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14
DB      462  GAGATGACACTGAAATTTGGTCAGCACGTAATGACAGTGTC 503

RESULT 7
AU279581
LOCUS   AU279581 CHONS2 Homo sapiens cDNA clone CHONS2000511 5', mRNA
DEFINITION
ACCESSION AU279581
VERSION   AU279581.1 GI:28298808
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS  Imabayashi H., Mori T., Gojo S., Kiyono T., Sugiyama T., Irie R.,
          Isogai T., Hata J., Tomoya Y. and Umezawa A.
TITLE    Redifferentiation of dedifferentiated chondrocytes and
          chondrogenesis of human bone marrow stromal cells via chondrosphere
          formation with expression profiling by large-scale cDNA analysis
JOURNAL  Exp. Cell Res. 288 (1), 35-50 (2003)
PUBMED   12878157
COMMENT  Contact: Takao Isogai
          Genomics Laboratory
          Helix Research Institute
          1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
          Tel: 81-438-52-3975
          Fax: 81-438-52-3986
          Email: genomics@hri.co.jp
          HRI human cDNA Project, Sugiyama T.; Wakamatsu A.; Irie R.;
          Umezawa A.; Fukuma M.; Kusakari S.; Hata J.; Ishii S.; Yamamoto J.;
          Isono Y.; Saito K.; Nakamura Y.; Masuho Y.; Nagai K.; Isogai T.
          HRI human cDNA project; cDNA library construction & 5'-end one
          pass sequencing; Helix Research Institute.
          Location/Qualifiers
FEATURES
    source
    1..538
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CHONS2000511"
    /cell_type="chondrocytes"
    /clone_lib="CHONS2"
    /note="Vector: pME18SFL3"

ORIGIN
Alignment Scores:
Pred. No.:      0.336      Length:      538
Score:          61.00      Matches:    11
Percent Similarity: 85.71%      Conservative: 1
Best Local Similarity: 78.57%      Mismatches: 2
Query Match:    75.31%      Indels:    0
Gap:            1          Gaps:      0

AAH26175_COPY_50_65 (1-16) x AU279581 (1-538)

QY      1  GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14
DB      416  GAGATGACACTGAAATTTGGTCAGCACGTAATGACAGTGTC 457

RESULT 8
AJ450203
LOCUS   AJ450203 riken1 Gallus gallus cDNA clone 24h8r1, mRNA
DEFINITION
ACCESSION AJ450203
VERSION   AJ450203.1 GI:20217424
KEYWORDS EST.
SOURCE   Gallus gallus (chicken)
ORGANISM
REFERENCE
AUTHORS  Suzuki Y., Yamashita R., Shirota M., Sakakibara Y., Chiba J.,
          Mizushima-Sugano J., Nakai K. and Sugano S.
TITLE    Sequence comparison of human and mouse genes reveals a homologous
          block structure in the promoter regions
JOURNAL  Genome Res. 14 (9), 1711-1718 (2004)
PUBMED   15342556
COMMENT  Contact: Yutaka Suzuki
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: yusuzuki@ims.u-tokyo.ac.jp.
          Location/Qualifiers
FEATURES
    source
    1..576
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="HRT00263"
    /tissue_type="heart"
    /clone_lib="Sugano cDNA library, heart"

ORIGIN
Alignment Scores:

```

```

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 571)
Buerstedde, J.M.
Gallus gallus bursal lymphocyte EST
Unpublished (2002)
Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
Location/Qualifiers
1..571
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="24h8r1"
/cell_type="bursal lymphocyte"
/dev_stages="2-3 weeks old"
/clone_lib="riken1"
/note="CB inbred strain"

ORIGIN
Alignment Scores:
Pred. No.:      0.359      Length:      571
Score:          61.00      Matches:    11
Percent Similarity: 85.71%      Conservative: 1
Best Local Similarity: 78.57%      Mismatches: 2
Query Match:    75.31%      Indels:    0
Gap:            1          Gaps:      0

AAH26175_COPY_50_65 (1-16) x AJ450203 (1-571)

QY      1  GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14
DB      495  GAGATGACACTGAAATTTGGTCGCTGTAATGATAGTGTC 536

RESULT 9
BP257160
LOCUS   BP257160 Sugano cDNA library, heart Homo sapiens cDNA clone
DEFINITION HRT00263, mRNA sequence.
ACCESSION BP257160
VERSION   BP257160.1 GI:52172390
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.
1 (bases 1 to 576)
Suzuki Y., Yamashita R., Shirota M., Sakakibara Y., Chiba J.,
Mizushima-Sugano J., Nakai K. and Sugano S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL  Genome Res. 14 (9), 1711-1718 (2004)
PUBMED   15342556
COMMENT  Contact: Yutaka Suzuki
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: yusuzuki@ims.u-tokyo.ac.jp.
          Location/Qualifiers
FEATURES
    source
    1..576
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="HRT00263"
    /tissue_type="heart"
    /clone_lib="Sugano cDNA library, heart"

ORIGIN
Alignment Scores:

```

```

Pred. No.: 0.363 Length: 576
Score: 61.00 Matches: 11
Percent Similarity: 85.71% Conservative: 1
Best Local Similarity: 78.57% Mismatches: 2
Query Match: 75.31% Indels: 0
DB: 3 Gaps: 0

AAH26175_COPY_50_65 (1-16) x BP257160 (1-576)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAanSerLysIle 14
   |||||
Db 410 GAGATGACACTGAAATTTGGTCCAGCAGCTAATGACAGTGTC 451

RESULT 10
LOCUS CN367889 577 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000424183436 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN367889
VERSION CN367889.1 GI:47367823
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominiidae; Homo.
REFERENCE 1 (bases 1 to 577)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
          Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
          Lebkowski,J and Stanton,L.W.
TITLE Transcriptome characterization elucidates signaling networks that
          control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
          Regenerative Medicine
          Geron Corporation
          230 Constitution Drive, Menlo Park, CA 94025, USA
          Tel: 650 473 8658
          Fax: 650 473 7760
          Email: rbrandenberger@geron.com
          Insert Length: 577 Std Error: 0.00.
FEATURES
          source
            1..577
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /tissue_type="embryonic stem cells, cell lines H1, H7, and
              H9"
              /clone_lib="GRN ES"
              /note="oligo dt primed, full-length enriched cDNA library
              from undifferentiated hES cell lines H1 (p32), H7 (p29),
              and H9 (p26) maintained in feeder-free conditions"

ALIGNMENT Scores:
Pred. No.: 0.364 Length: 577
Score: 61.00 Matches: 11
Percent Similarity: 85.71% Conservative: 1
Best Local Similarity: 78.57% Mismatches: 2
Query Match: 75.31% Indels: 0
DB: 7 Gaps: 0

AAH26175_COPY_50_65 (1-16) x CN367889 (1-577)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAanSerLysIle 14
   |||||
Db 289 GAGATGACACTGAAATTTGGTCCAGCAGCTAATGACAGTGTC 330

RESULT 11
LOCUS BP201994 582 bp mRNA linear EST 14-SEP-2004
DEFINITION BP201994 Sugano cDNA library, amygdala Homo sapiens cDNA clone
          AMR07079, mRNA sequence.

```

```

ACCESSION BP201994
VERSION BP201994.1 GI:52052519
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominiidae; Homo.
REFERENCE 1 (bases 1 to 582)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
          Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
          block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES
          source
            1..582
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="AMR07079"
              /tissue_type="amygdala"
              /clone_lib="Sugano cDNA library, amygdala"

ORIGIN
Alignment Scores:
Pred. No.: 0.367 Length: 582
Score: 61.00 Matches: 11
Percent Similarity: 85.71% Conservative: 1
Best Local Similarity: 78.57% Mismatches: 2
Query Match: 75.31% Indels: 0
DB: 3 Gaps: 0

AAH26175_COPY_50_65 (1-16) x BP201994 (1-582)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAanSerLysIle 14
   |||||
Db 433 GAGATGACACTGAAATTTGGTCCAGCAGCTAATGACAGTGTC 474

RESULT 12
LOCUS BQ037883 606 bp mRNA linear EST 01-MAY-2002
DEFINITION pgnic.pk007.j22 normalized chicken lymphoid cDNA library Gallus
          gallus cDNA clone pgnic.pk007.j22 5', similar to ref|NP_001871.1
          (NM_001880) activating transcription factor 2; CAMP-response
          element-binding protein-2; CAMP responsive element binding protein
          2 [Homo sapiens] pir|S05380 transcription factor ATF2 - human
          emb|CAA33886.1| (X15875) CAMP response, mRNA sequence.
ACCESSION BQ037883
VERSION BQ037883.2 GI:20382615
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
          Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 606)
AUTHORS Morgan,R.W. and Burnside,J.
TITLE Chicken lymphoid ESTs
JOURNAL Unpublished (2001)
COMMENT On Mar 27, 2002 this sequence version replaced gi:19771423.
          Contact: Joan Burnside
          Molecular Endocrinology
          University of Delaware
          40 Townsend Hall, Newark, DE 19717, USA
          Tel: 302 831-1345
          Fax: 302-831-3411
          Email: joan@Udel.Edu, www.chickest.udel.edu.

```

```

FEATURES
source
Location/Qualifiers
1..606
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="pgnic.pk007.j22"
/sex="Male and Female"
/tissue_type="thymus, bursa, spleen, PBL, bone marrow"
/lab_host="E.coli BMDH108"
/clone_lib="normalized chicken lymphoid cDNA library"
/notes="Vector: pCMVSPORT 6"

ORIGIN
Alignment Scores:
Pred. No.: 0.385 Length: 606
Score: 61.00 Matches: 11
Percent Similarity: 85.71% Conservative: 1
Best Local Similarity: 78.57% Mismatches: 2
Query Match: 75.31% Indels: 0
DB: 3 Gaps: 0

AAH26175_COPY_50_65 (1-16) x BQ037883 (1-606)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14
|||||
DB 268 GAGATGACACTGAATTTGGTCCGGCTCTGTAATGATAGTGTC 309

RESULT 13
BU662818 613 bp mRNA linear EST 30-SEP-2002
LOCUS
DEFINITION
c191g01.z1 Hembase; Erythroid Precursor Cells (LCB:cl library) Homo
sapiens cDNA clone c191g01 5', mRNA sequence.
ACCESSION
BU662818
VERSION
BU662818.1 GI:23375003
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 613)
Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
Gene Expression in Human Erythroid Precursor Cells
Unpublished (2002)
Contact: Jeffrey L. Miller
Laboratory of Chemical Biology
National Institute of Diabetes and Digestive and Kidney Diseases
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
20892, USA
Tel: 301 402 2373
Fax: 301 435 5148
Email: jm7f@nih.gov
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the
Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
analyses by National Institutes of Health Intramural Sequencing
Center (NISC). More information available at:
http://hembase.nidk.nih.gov
Plate: 91 row: g column: 01
Seq primer: 5' lambda-Triplex2 Sequencing Primer.

FEATURES
source
Location/Qualifiers
1..613
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="c191g01"
/sex="unknown"
/tissue_type="blood"
/cell_type="Erythroid Precursor Cells"
/cell_line="Primary Culture of Peripheral Blood
Mononuclear Cells"
/dev_stage="Precursor erythroblasts; GPA++"
/lab_host="DH5alpha"
/clone_lib="Hembase; Erythroid Precursor Cells (LCB:cl
library)"
/notes="Organ: blood; Vector: pTriplex2; Site 1: SfiI;
Site 2: SfiI; A complementary DNA (cDNA) library from
human erythroid precursor cells was constructed using
SMART PCR (polymerase chain reaction) cDNA Library
Construction Kit (Clontech, Palo Alto, CA) according to
the manufacturer's directions, but with slight
modifications. Briefly, reverse transcription was
performed in the presence of 1 umol/L peptide nucleic acid
(PNA) oligos
(N-terminal)-biotin-GTC-CAC-CCG-AGG-CTT-G-(C-terminal) and
(N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal)
. Synthesized cDNA was digested with SfiI and
size-selected on a 1% agarose gel (>800bp). Large-scale
sequencing of the library was performed by the NIH
Intramural Sequencing Center (NISC,
Http://www.nisc.nih.gov/)."

ORIGIN
Alignment Scores:
Pred. No.: 0.39 Length: 613
Score: 61.00 Matches: 11
Percent Similarity: 85.71% Conservative: 1
Best Local Similarity: 78.57% Mismatches: 2
Query Match: 75.31% Indels: 0
DB: 5 Gaps: 0

AAH26175_COPY_50_65 (1-16) x BU662818 (1-613)

QY 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14
|||||
DB 165 GAGATGACACTGAATTTGGTCCAGCAGCTAATGACAGTGTC 206

RESULT 14
CN367892 621 bp mRNA linear EST 16-MAY-2004
LOCUS
DEFINITION
17000532626080 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION
CN367892
VERSION
CN367892.1 GI:47367826
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 621)
Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,X., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J. and Stanton,L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
15146197
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert length: 621 Std Error: 0.00.

FEATURES
source
Location/Qualifiers
1..621
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN EB"
/notes="oligo dT primed, full-length enriched cDNA library
from embryoid body outgrowths derived from HES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

```

Search completed: January 12, 2006, 20:45:39  
Job time : 2611 secs

```

ORIGIN
Alignment Scores:
Pred. No.:      0.395      Length:      621
Score:          61.00      Matches:    11
Percent Similarity: 85.71%  Conservative: 1
Best Local Similarity: 78.57% Mismatches: 2
Query Match:      75.31%   Indels:      0
DB:               7       Gaps:        0

AAH26175_COPY_50_65 (1-16) x CN367892 (1-621)

QY      1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14
      |||||||
DB      401 GAGATGACACTGAATTTGGTCCAGCACGTAATGACAGTGTC 442
      |||||||

RESULT 15
AL119505
LOCUS      635 bp      mRNA      linear      EST 04-SEP-2003
DEFINITION DKFZp761O191_r1_761 (synonym: hamy2) Homo sapiens cDNA clone
ACCESSION  AL119505
VERSION    AL119505.1 GI:5925404
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominoidea; Homo.
REFERENCE  1 (bases 1 to 635)
AUTHORS   Ottenwaelder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and
            Wiemann, S.
TITLE     EST (Ottenwaelder, et al.)
JOURNAL   Unpublished (1999)
COMMENT   Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by MediGenomix (Martinsried/Germany) within the cDNA
            sequencing consortium of the German Genome Project. No s1 sequence
            available.
            This clone (DKFZp761O191) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
            source
            1..635
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="DKFZp761O191"
               /tissue_type="amygdala"
               /dev_stage="adult"
               /lab_host="DH10B"
               /clone_lib="761 (synonym: hamy2)"
               /note="vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN
Alignment Scores:
Pred. No.:      0.406      Length:      635
Score:          61.00      Matches:    11
Percent Similarity: 85.71%  Conservative: 1
Best Local Similarity: 78.57% Mismatches: 2
Query Match:      75.31%   Indels:      0
DB:               1       Gaps:        0

AAH26175_COPY_50_65 (1-16) x AL119505 (1-635)

QY      1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14
      |||||||
DB      433 GAGATGACACTGAATTTGGTCCAGCACGTAATGACAGTGTC 474
      |||||||

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2006, 18:44:35 ; Search time 141 Seconds  
(without alignments)  
201.709 Million cell updates/sec

Title: AAH26175\_COPY\_50\_65  
Perfect score: 81  
Sequence: 1 emtlkfgparksiee 16

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1303057 segs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlh  
-Q=/cgn2\_1/USPTO.spool/RAWLINGS10076905/runat\_10012006.144825\_5355/app\_query.fasta\_1.199  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=RAWLINGS10076905 @CGN\_1\_1\_193 @runat\_10012006.144825\_5355 -NCPUS=6  
-ICPU=3 -NO\_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/PTCUS COMB.seq:\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq:\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq:\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	75.3	1621	3	US-09-949-016-5410
2	61	75.3	65966	3	US-09-949-016-17152
3	55	67.9	299	3	US-09-513-999C-2946
4	45	55.6	29604	3	US-08-781-891-207
5	45	55.6	29604	3	US-09-618-166-207
6	45	55.6	1664976	3	US-08-916-421B-1
C 7	45	55.6	1664976	3	US-09-692-570-1
C 8	43	53.1	11936	3	US-09-949-016-17428
9	43	53.1	4403765	3	US-09-103-840A-2

10	43	53.1	4411529	3	US-09-103-840A-1	Sequence 1, Appli
C 11	42	51.9	237	3	US-09-134-000C-2916	Sequence 2916, Ap
C 12	42	51.9	450	3	US-09-252-991A-664	Sequence 664, App
C 13	42	51.9	778	3	US-09-270-767-6315	Sequence 6315, Ap
C 14	42	51.9	778	3	US-09-270-767-6315	Sequence 21597, A
C 15	42	51.9	1407	3	US-09-252-991A-637	Sequence 637, App
C 16	42	51.9	1728	3	US-09-252-991A-616	Sequence 616, App
C 17	41	50.6	601	3	US-09-949-002-1683	Sequence 1683, Ap
C 18	41	50.6	601	3	US-09-949-002-9306	Sequence 9306, Ap
C 19	41	50.6	2562	3	US-09-620-312D-264	Sequence 264, App
C 20	41	50.6	11048	3	US-09-949-016-17018	Sequence 17018, A
C 21	41	50.6	16039	3	US-09-949-002-812	Sequence 812, App
C 22	41	50.6	26760	3	US-09-949-016-15894	Sequence 15894, A
C 23	41	50.6	47787	3	US-09-949-016-11969	Sequence 11969, A
C 24	41	50.6	63658	3	US-09-949-016-13238	Sequence 13238, A
C 25	41	50.6	64489	3	US-09-949-016-11766	Sequence 11766, A
C 26	41	50.6	98962	3	US-09-949-016-14133	Sequence 14133, A
C 27	41	50.6	102884	3	US-09-949-016-17100	Sequence 17100, A
C 28	41	50.6	536165	3	US-09-214-808-1	Sequence 1, Appli
C 29	40	49.4	591	3	US-09-533-559-3318	Sequence 3318, Ap
C 30	40	49.4	601	3	US-09-949-016-24450	Sequence 24450, A
C 31	40	49.4	601	3	US-09-949-016-198833	Sequence 198833, Ap
C 32	40	49.4	726	3	US-09-248-796A-9963	Sequence 9963, Ap
C 33	40	49.4	1212	3	US-09-248-796A-136	Sequence 136, App
C 34	40	49.4	1314	3	US-09-248-796A-8070	Sequence 8070, Ap
C 35	40	49.4	1747	4	US-09-605-703B-1265	Sequence 1265, Ap
C 36	40	49.4	1899	4	US-09-605-703B-1263	Sequence 1263, Ap
C 37	40	49.4	7479	3	US-09-991-258-2	Sequence 2, Appli
C 38	40	49.4	11459	3	US-09-454-721A-3	Sequence 3, Appli
C 39	40	49.4	12379	3	US-09-991-258-14	Sequence 14, Appli
C 40	40	49.4	12523	3	US-09-991-258-1	Sequence 1, Appli
C 41	40	49.4	13584	3	US-09-991-258-17	Sequence 17, Appli
C 42	40	49.4	373182	3	US-09-949-016-17171	Sequence 17371, A
C 43	40	49.4	373894	3	US-09-949-016-12062	Sequence 12062, A
C 44	39.5	48.8	95255	3	US-09-949-016-17067	Sequence 17067, A
C 45	39	48.1	474	3	US-09-902-540-6590	Sequence 6590, A

#### ALIGNMENTS

RESULT 1

US-09-949-016-5410  
; Sequence 5410, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5410

; LENGTH: 1621

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-5410

Alignment Scores:

Score: 0.0134 Length: 1621

Percent Similarity: 61.00 Matches: 11

Best Local Similarity: 85.71% Conservative: 1

Query Match: 78.57% Mismatches: 2

Indels: 0

DB: 3 Gaps: 0







```
NAME/KEY: misc feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
```

```

/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1119881)..(1119881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1130881)..(1130881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1310988)..(1310988)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1313224)..(1313224)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1349473)..(1349473)
/ OTHER INFORMATION: n equals a, t, c, or g

```

Alignment Scores:	1.27e+05	1664976
Pred. No.:		
Score:	45.00	9
Percent Similarity:	68.75%	Conservative: 2
Best Local Similarity:	65.25%	Mismatches: 5
Query Match:	55.56%	Indels: 0
DB:	3	Gaps: 0

AAH26175 COPY 50 65 (1-16) X US-09-692-570-1 (1-1664976)

1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIleGluGlu 16  
 1471514 GAATCAAAATTAATAATCAGACCGTTTCGGAATGGAAAGATAGAGAA 1471467

```

RESULT 8
US-09-949-016-17428
; Sequence 17428, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17428
; LENGTH: 11936
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17428

```

Alignment Scores:		
Pred. No.:	733	Length: 11936
Score:	43.00	Matches: 8
Percent Similarity:	78.57%	Conservative: 3
Best Local Similarity:	57.14%	Mismatches: 3
Query Match:	53.09%	Indels: 0
DB:	3	Gaps: 0

AAH26175 COPY 50 65 (1-16) X-11S-09-949-016-17428 (1-11936)

Qy 1 GluMetThrLeuLysPheGlyProAlaArgAsnSerLysIle 14

9063 GAGGTGACACTTGTGAGTTGGGGCCTCAAGGACAAGTAAGATG 9104

```

RESULT 9
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24365-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

```

Alignment Scores:	9.02e+05	Length:	4403765
Pred. No.:		Matches:	8
Score:	43.00	Conservative:	1
Percent Similarity:	80.00%	Mismatches:	1
Best Local Similarity:	90.00%	Indels:	0
Query Match:	53.09%	Gaps:	0
DB:	3		

22H26175 COPY 50 65 (1-16) x US-09-103-840A-2 (1-4403765)

Qy 4 LeuLysPheGlyProAlaArgAsnSerLys 13  
::: |||||  
nb 3531174 GTGAAATAGGCCCGGCCCGAAATTCGAAA 3531203

```

RESULT 10
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 441529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 9.04e+05 441529
Score: 43.00 8
Percent Similarity: 90.00% 1
Best Local Similarity: 80.00% 1
Query Match: 53.00% 0
Gaps: 3 0
DB:

```

AAH26175\_COPY\_50\_65 (1-16) x US-09-103-840A-1 (1-4411529)

QY 4 LeuLysPheGlyProAlaArgAenSerLys 13  
Db 3533592 GTGAAATAGGCGCCGCGGAATCGAA 3533621

RESULT 11  
US-09-134-000C-2916/c  
; Sequence 2916, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2916  
; LENGTH: 237  
; TYPE: DNA  
; ORGANISM: Enterococcus faecalis

US-09-134-000C-2916  
Alignment Scores:  
Pred. No.: 9.12 Length: 237  
Score: 42.00 Matches: 8  
Percent Similarity: 73.33% Conservative: 3  
Best Local Similarity: 53.33% Mismatches: 4  
Query Match: 51.85% Indels: 0  
DB: 3 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x US-09-134-000C-2916 (1-237)

QY 2 MetThrLeuLysPheGlyProAlaArgAenSerLysIleGluGlu 16  
Db 114 ATGTCACACCCATGGACCATACCTTAATACAAAATAGAGAA 70

RESULT 12  
US-09-252-991A-664  
; Sequence 664, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 664  
; LENGTH: 450  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-664  
Alignment Scores:  
Pred. No.: 20.2 Length: 450  
Score: 42.00 Matches: 8  
Percent Similarity: 78.57% Conservative: 3  
Best Local Similarity: 57.14% Mismatches: 3  
Query Match: 51.85% Indels: 0  
DB: 3 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x US-09-252-991A-664 (1-450)

QY 3 ThrLeuLysPheGlyProAlaArgAenSerLysIleGluGlu 16  
Db 42 ACCCTGCAATTCGCGCCGACCTTCAACGCCAGCCGCGAGAG 83

RESULT 13  
US-09-270-767-6315/c  
; Sequence 6315, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6315  
; LENGTH: 778  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster

US-09-270-767-6315  
Alignment Scores:  
Pred. No.: 39.8 Length: 778  
Score: 42.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 51.85% Indels: 0  
DB: 3 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x US-09-270-767-6315 (1-778)

QY 3 ThrLeuLysPheGlyProAlaArgAen 11  
Db 339 ACACAAAATTTGGACCACGACAGAAAT 313

RESULT 14  
US-09-270-767-21597/c  
; Sequence 21597, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 21597  
; LENGTH: 778  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster

US-09-270-767-21597  
Alignment Scores:  
Pred. No.: 39.8 Length: 778  
Score: 42.00 Matches: 8  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 1  
Query Match: 51.85% Indels: 0  
DB: 3 Gaps: 0

AAH26175\_COPY\_50\_65 (1-16) x US-09-270-767-21597 (1-778)

QY 3 ThrLeuLysPheGlyProAlaArgAen 11  
Db 339 ACACAAAATTTGGACCACGACAGAAAT 313

RESULT 15  
US-09-252-991A-637  
; Sequence 637, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:

Fri Jan 13 09:47:01 2006

```

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 637
; LENGTH: 1407
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-637

```

```

Alignment Scores:
Pred. No.:      82.9      Length:      1407
Score:          42.00     Matches:      8
Percent Similarity: 78.57% Conservative: 3
Best Local Similarity: 57.14% Mismatches: 3
Query Match:      51.85% Indels:      0
DB:               3      Gaps:      0

AAH26175_COPY_50_65 (1-16) x US-09-252-991A-637 (1-1407)
QY      3  ThrLeuLysPheGlyProAlaArgAsnSerLysIleGlu 16
Db      1315 ACCCTGCAATTGGCCCGACCTTCACGCCAGCGCGGAGAG 1356

```

Search completed: January 12, 2006, 21:11:26  
Job time : 1545 secs



This Page is inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record

## BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☒ OTHER: Side texts

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images  
problems checked, please do not report the  
problems to the IFW Image Problem Mailbox**